University of Bath

PHD

Condition monitoring of transformers based on non-invasive measurements and characterisation of partial discharges

Babnik, Tadeja

Award date:
2005

Awarding institution:
University of Bath

Link to publication

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (https://creativecommons.org/licenses/by-nc-nd/4.0/). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

Take down policy
If you consider content within Bath's Research Portal to be in breach of UK law, please contact: openaccess@bath.ac.uk with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

Download date: 05. May. 2023

UNIVERSITY OF
BATH

CONDITION MONITORING OF TRANSFORMERS BASED ON NON-INVASIVE MEASUREMENTS AND CHARACTERISATION OF PARTIAL DISCHARGES

Submitted by Tadeja Babnik
for the degree of PhD
of the University of Bath

2005

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with its author. This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the prior written consent of the author.

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation.

Tadeja Babnik

UMI Number: U201066

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U201066
Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

UNIVERSITY OF BATH
LIBRARY
12 JAN 2006
PhD

To my Mum, Sis, Bojan and little Tina

Acknowledgements

I would like to acknowledge my supervisors Prof. R.K. Aggarwal and Prof. P.J. Moore for their support, guidance and advices during my research activities and the preparation of this thesis.

This work was financially supported by National Grid Company. They also arranged the testing facilities throughout the on-site tests at the Neilston Substation in Scotland and at the Northfleet substation in England. I would like to thank for the assistance and expertise of their staff, especially Paul Jarman, Frank Bond and Denis Procter. I would also like to thank the team at the Melksham substation, particularly Mr. Tim Adams for making the measurements in the substation possible.

Dr. Zhongdong Wang gave me an opportunity to make measurements at the High voltage laboratory at UMIST. My special thanks go to her.

I wish to thank Prof. W.R. Rutgers from KEMA for many useful and extensive discussions.

I am also thankful to my friends and colleagues in Bath who made my stay with them pleasant and unforgettable.

Finally, I would like to express my deepest appreciation to my family for their unreserved support, trust and patience every day. I wish to thank my sister for all phone calls when I had hard time and difficulties in England. Bojan, thanks to you I know the weight of real solutions in the power engineering.

Abstract

This thesis describes the investigation and development of a non-invasive approach for detection and interpretation of partial discharges dedicated to power transformers. The presented method is based on remote radiometric measurements. The advantage of the method is that it requires no galvanic connections to the transformer, thus allowing measurements when the transformer is in operation. Since fingerprints of remote radiometric measurements are not available, the formation of clusters with similar features obtained from captured partial discharge (PD) data was crucial and this included the implementation of a number of methods for clustering captured data.

Apart from a description of the fundamental design of the measuring system, this thesis focuses on the PD signals analysis, based on on-site PD measurements captured by antennas and the KEMA probe. In view of the fact that there were thousands of signals captured in the substation and each signal was represented by 5000 samples, the dimensionality reduction was essential; for this purpose, principal component analysis technique was employed. Through the application of the proposed methods to the PD data, it is demonstrated that self-organising map (SOM) and cluster analysis (CA) are able to form groups encompassing similarities. Several similarity measurers and linkage algorithm were applied to determine the most appropriate method. For clustering PD signals by CA, it is shown that the Ward's linkage algorithm is the most appropriate, since it enables distinct groups to be formed with a relatively high number of signals within each group and importantly, it does not seek individuals. Moreover, a clear separation of PDs emanating from a transformer and discharges emanating from the surrounding is achieved. Tests in the high voltage Laboratory proven that based on the frequency spectra of the signals, the characteristics of discharges can be identified.

Radio metric measurements have proved to be a promising tool in detecting and identifying different discharges. For pattern recognition purposes, it is important to create a data base of many practically encountered different signals and hence more tests need to be performed in order to gain a better understanding of the PD pulses.

Contents

Acknowledgements i

Abstract ii

Contents iii

List of Abbreviations x

List of Symbols xii

Index to Figures xv

Index to Tables xxi

1. INTRODUCTION 1

1.1. Overview 1

1.2. Objective of the Project 4

1.3. Scope of the thesis 4

2. COMPONENTS AND PROPERTIES OF A TRANSFORMER INSULATION SYSTEM 6

2.1. Transformer construction 6

2.2. Transformer Insulation 7

2.2.1. Solid insulation 7

2.2.2. Mineral oil 8

2.2.3. Oil impregnated paper insulation 8

2.3. Insulation Degradation and Ageing 9

2.3.1. Degradation of oil-paper insulation systems 9

3. ASSESSMENT OF POWER TRANSFORMER CONDITION 11

3.1. Introduction 11

3.2. Electrical methods 11

3.2.1. Leakage inductance measurement 11

3.2.2. Measurement of magnetizing currents 12

3.2.3. Insulation resistance testing and polarization index 12

3.2.4. Recovery (return) voltage measurement - RVM 14

3.2.5. Polarisation and depolarisation currents (PDC) 15

3.2.6. Loss-factor (dissipation factor, tan δ) and capacitance measurement 15
3.2.7. Frequency response analysis (FRA) 16
3.2.8. Partial discharge (PD) measurement 16
3.3. Chemical methods 17
3.3.1. Oil analysis 17
3.3.2. Dissolved gas analysis (DGA) 17
3.3.3. Degree of polymerization (DP) 18
3.3.4. High performance liquid chromatography (HPLC) 19
3.4. Physical methods 19
3.4.1. Radio interference 19
3.4.2. Infrared emission 20
3.4.3. Acoustic emission 20
3.4.4. Vibration 20
3.5. Conclusion on diagnostic methods 21
3.6. Power transformer monitoring 21
3.6.1. Introduction 21
3.6.2. Gas-in-oil 22
3.6.3. Partial discharges (PDs) 22
3.6.4. Ultrasonic and vibration detection 23
3.6.5. Temperature 23
3.6.6. Water in oil or cellulose 24
3.6.7. On load tap changer 24
3.6.8. Integration 25
3.6.9. Summary on the transformer monitoring equipment 25
4. PARTIAL DISCHARGES 27
4.1. Introduction 27
4.2. Occurrence of discharges [56] 28
4.2.1. Inception of internal discharges 28
4.2.2. Inception of surface discharges 29
4.2.3. Inception of corona discharges 29
4.2.4. Discharges by electrical treeing 29
4.3. Recurrence of discharges 29
4.4. PD sources in transformers 32

5. PD MEASUREMENTS AS A DIAGNOSTIC TOOL FOR POWER TRANSFORMERS 34

5.1. Introduction 34

5.2. Definitions of Quantities Related to Partial Discharges [12] 35

5.3. PD activity detection 35

5.3.1. Acoustic detection of PD activity 35

5.3.2. Optical detection of PD activity 36

5.3.3. Chemical detection of PD activity 37

5.3.4. Electrical measurement of PD activity 37

5.4. PD Pattern Analysis Techniques 39

5.4.1. Phase related PD recognition 40

5.4.1.1. Phase resolved partial discharge (PRPD) pattern 41

5.4.1.2. Δu pattern 42

5.4.1.3. Δu/Δφ Pattern (pulse sequence pattern) 42

5.4.2. Time resolved recognition 43

5.5. Classification and decision 44

6. POWER TRANSFORMER NON-INVASIVE PD MEASUREMENTS – FIRST TRIAL 45

6.1. Introduction 45

6.2. Test Setup 45

6.3. Partial Discharge Acquisition System 46

6.4. Results 47

6.5. Conclusion and directions 54

7. WIDE BAND MEASUREMENTS 55

7.1. Introduction 55

7.2. Set Up Of The Detection System 55

7.3. Typical measured signals 57

7.3.1. Typical signals obtained by Yagi 1 58

7.3.2. Typical signals obtained by H1 61

7.4. Conclusions on wide band measurements 65

8. DATA MINING 66

8.1. Introduction **66**

8.2. Data Mining Process **67**

8.3. Data Mining Tasks **68**

8.4. Summarization **68**

8.4.1. Clustering 69

8.4.2. Classification 69

8.4.3. Regression 69

8.4.4. Dependency Modelling 69

8.4.5. Deviation Detection 69

8.4.6. Temporal Problems 70

8.4.7. Causation Modelling 70

8.5. Data Mining Techniques **70**

8.5.1. Tree and Rule Induction 70

8.5.2. Association Rules 70

8.5.3. Clustering Methods 71

8.5.4. Artificial Neural Networks 71

8.5.5. Statistical Techniques 71

8.5.6. Visualization Techniques 71

8.5.7. Additional Tools 71

8.5.8. Common Features 72

8.6. DM Operational Aspects **72**

8.6.1. The Right Tools 72

8.6.2. The Right Data 73

8.6.3. The Right People 73

8.6.4. The Right Application 73

8.7. Dimensionality reduction **73**

8.7.1. Introduction 73

8.7.2. Feature Extraction 74

8.7.3. Principal Component Analysis 75

8.7.3.1. Introduction *75*

8.7.3.2. Principal Components *76*

8.7.3.3. Simple example illustrating PCA *77*

8.7.4. Method of moments 78

8.8. Self-Organising Map **79**

8.8.1. The structure 80

8.8.1.1. Initialisation *81*
8.8.1.2. The training process *81*
8.8.1.3. Visualization *83*
8.8.1.4. The configuration of SOM *84*
8.8.2. Software Tools of SOM 85
8.8.3. An Example of SOM application 86
8.9. Cluster analysis [99]- [101] 88
8.9.1. Introduction 88
8.9.2. Basic principles 88
8.9.3. Objects selection 89
8.9.4. Interval-scaled variables (Standardizing the data) 89
8.9.5. Similarity and dissimilarity measures 90
8.9.5.1. Distance measures *90*
8.9.5.2. Association coefficients *91*
8.9.5.3. Probabilistic similarity coefficients *91*
8.9.6. CA techniques 91
8.9.7. Hierarchical methods 92
8.9.8. Assessment of solution 96
8.9.9. An example of hierarchical clustering 97
8.10. Summary 99
9. DATA MINING ON THE PD DATA USING THE SOM 100
9.1. Introduction 100
9.2. A Feasibility Study Based on the Proposed Approach 100
9.3. Analysis on revealed features of the PD data 103
10. DATA MINING ON THE PD DATA USING THE CA 118
10.1. Introduction 118
10.2. A Feasibility Study based on the Proposed Approach 118
10.2.1. Results of CA 120
10.3. Summary on data mining on the PD data using the CA 129
11. A COMPARISON OF RESULTS OBTAINED BY SOM AND CA 131
11.1. Summary of observations 142
12. RADIO FREQUENCY MEASUREMENT OF DIFFERENT DISCHARGE SOURCES AT THE UMIST LABORATORY 144

12.1. Introduction **144**

12.2. Models of defects **144**

12.3. Measurement configuration **146**

12.4. Measurement results and analysis **146**

12.5. Summary **149**

13. METHOD OF MOMENTS APPLIED TO PD DATA **150**

13.1. Introduction **150**

13.2. PD measurements captured at the Northfleet substation in time-frequency plane **150**

13.3. Comparison of PD measurements captured at the Northfleet and the Melksham substation in time-frequency plane **151**

13.4. Measurements obtained at the UMIST Laboratory in time-frequency plane **152**

13.5. Summary **153**

14. CONCLUSIONS AND FUTURE WORK **154**

14.1. Introduction **154**

14.2. Review of Major Studies and Achievements Attained in This Research Project **155**

14.2.1. A design of non-invasive PD measurements dedicating to power transformers 155

14.2.2. Dimensionality reduction of captured data 156

14.2.3. Application of SOM on PD data 156

14.2.4. Application of CA on PD data 156

14.2.5. Analysis of signals originating from different defect models in the HV Laboratory 157

14.2.6. Method of moments applied to the PD data 157

14.2.7. Final conclusions 157

14.3. Future Work **158**

15. REFERENCES **160**

16. APPENDIX **168**

16.1. Results of the measurements obtained by Yagi 2 ..... **168**

16.2. Results of the measurements obtained by Yagi 3 ..... **171**

16.3. Results of the measurements obtained by H2 ..... **173**

16.4. Results of the measurements obtained by H3 175

16.5. Results of the measurements obtained by H4 177

Related publications 179

List of Abbreviations

AC	Alternating Current
ANN	Artificial Neural Network
BGN	Background noise
BMU	Best matching unit
CA	Cluster Analysis
CIGRE	Conseil International des Grands Réseaux Électriques
DC	Direct Current
DG	Degree of Polymerization
DGA	Dissolved Gas Analysis
DM	Data Mining
EM	Electro magnetic
FP	Floating particle
FRA	Frequency Response Analysis
H1	Helical antenna placed at position 1
H2	Helical antenna placed at position 2
H3	Helical antenna placed at position 3
H4	Helical antenna placed at position 4
HPLC	High Performance Liquid Chromatography
HV	High Voltage
ICA	Independent Component Analysis
IEC	International Electrotechnical Commission
IR	Insulation Resistance
K-L	Karhunen-Loève
LTC	Load Tap Changer
LV	Low Voltage
MDS	Multidimensional Scaling
ML	Machine Learning
OLTC	On-load Tap Changer
PC	Principal Component
PCA	Principal Component Analysis
PD	Partial Discharge
PDC	Polarisation and Depolarisation Currents
PI	Polarization Index
PRPD	Phase resolved partial discharge

RFI	Radio Frequency Interference
RVM	Recovery Voltage Measurement
SGT	Super Grid Transformer
SOM	Self Organizing Map

List of Symbols

%	Percent
ϵ	Dielectric permittivity
δ	Loss factor
σ	Conductivity
Cu	Copper
Pb	Palladium
I_{P1}	Polarization index
$i_{15"}$	Current measured after 15 seconds (nA)
$i_{60"}$, $i_{1'}$	Current measured after 60 seconds (nA)
$i_{10'}$	Current measured after 10 minutes (nA)
$R_{10'}$	Insulation resistance (MΩ)
U_0	DC voltage (V)
$U_R(t)$	Recovery voltage (V)
t_c	Charging time
t_g	Short-circuit time
U_{Rmax}	Peak of the recovery voltage (V)
dU_R/dt	Initial rate of rise of the recovery voltage
H_2	Hydrogen
CH_4	Methane
C_2H_6	Ethane
C_2H_4	Ethylene
C_2H_2	Acetylene
CO_2	Carbon dioxide
CO	Carbon monoxide
V_c	Voltage across the void
V_a	Applied voltage
C_a	Capacitance of the sound part of the dielectric
C_b	Capacity of the dielectric in series with the cavity
C_c	Capacity of the cavity
V^+	Ignited voltage
q	Apparent charge
n	Pulse repetition rate
N	Pulse repetition frequency
I	Average discharge current
D	Quadratic rate

P	Discharge power
kHz	Kilo Hertz
f_1	Lower frequency
f_2	Upper frequency
μs	Micro second
T_r	Pulse resolution time
f_m	Midband frequency
Δf	Frequency bandwidth
Vi	Ignition voltage
ϕ	Phase angle
H_N	Pulse count distribution
H_{qn}	Mean pulse height distribution
H_{qmax}	Maximum pulse height distribution
$H_{\Sigma q}$	Average discharge current distribution
S_k	Skewness
K_u	Kurtosis
P_e	Number of peaks
mcc	Cross correlation factor
$\Delta u/\Delta\varphi$	Pulse sequence pattern
m	Slope
U_p	Test voltage
V	Voltage
V_r	Rated voltage
f	Frequency
S	Covariance matrix
λ	Eigenvalue
e	Eigenvector
σ_T	Equivalent time
σ_F	Equivalent frequency
$\mathcal{R}^n$	N-dimensional input space
n	Dimension of input space
m_i	Reference vector of neuron i
x	Input vector
c	Best matching unit
$h_{ci}(t)$	Neighbourhood kernel centred on c at t
$h(\|\|rc - ri\|\|, t)$	Neighbourhood function at t

$\alpha(t)$	Learning-rate function at t
num_neuron	Total number of neurons
dlen	Total number of input vectors
x-dim, *y*-dim	Map-sizes for x and y-dimension of SOM
ε_{AQ}	Average quantisation error
ε_T	Topographic error
d	Distance
S	Silhouette width

Index to Figures

Figure 3.1: Real time inputs to the monitoring system 26
Figure 4.1: Internal discharges 28
Figure 4.2: Surface discharges 28
Figure 4.3: Corona discharges 28
Figure 4.4: Discharges in electrical threes 28
Figure 4.5: Void in a solid dielectric and equivalent circuit 30
Figure 4.6: Sequence of cavity breakdown under AC voltage 31
Figure 4.7: A dielectric circuit for surface discharges 31
Figure 4.8: An analogue circuit for corona discharges 32
Figure 5.1: A general procedure of PD measurement and evaluation 34
Figure 5.2: Conventional PD measurement circuit 38
Figure 5.3: A PD pulse sequence and relevant parameters 40
Figure 5.4: Three stages of phase related recognition 41
Figure 5.5: φ-q-N pattern 42
Figure 5.6: Example of Δu pattern H(Δui+1/Δui) 42
Figure 5.7: Voltage gradient and voltage difference of two consecutive pulses 43
Figure 5.8: Voltage gradient and voltage difference of two consecutive pulses 43
Figure 5.9: Pulse parameters in the time domain 44
Figure 6.1: A diesel generator unit and a step up transformer 0.433/33 kV (left) and a 400/275/33 kV transformer (right) under test for PD-diagnosis 46
Figure 6.2: Ground plan of tested transformer unit with generator unit and step-up transformer and antenna locations. 46
Figure 6.3: Schematic representation of the PD measuring system: set-up of principal components 47
Figure 6.4: Captured PD signals and background noise by monopole antenna 48
Figure 6.5: Captured PD signals and background noise by Yagi antenna placed beside the cabin 48
Figure 6.6: Measured PDs signal by monopole antenna together with the voltage signal 49
Figure 6.7: PD measurement by monopole antenna, $V = 100\%V_r, f = 50$ Hz, location 1 50
Figure 6.8: PD measurement by helical antenna, $V = 100\%V_r, f = 50$ Hz, location 3 50
Figure 6.9: PD measurement by Yagi antenna in phase L3, vertical polarisation, $V = 100\%V_r, f = 50$ Hz, location 2 50
Figure 6.10: PD measurement by Yagi antenna in phase L3, horizontal polarisation, $V = 100\%V_r, f = 50$ Hz, location 2 50

Figure 6.11: PD measurement by Yagi antenna in phase L2, vertical polarisation, $V = 100\%V_r, f = 50$ Hz, location 2 50
Figure 6.12: PD measurement by Yagi antenna in phase L2, horizontal polarisation, $V = 100\%Vr, f = 50$ Hz, location 2 50
Figure 6.13: PD measurement by Yagi antenna beside cabin, vertical polarisation, $V = 100\%V_r, f = 50$ Hz, location 1 51
Figure 6.14: PD measurement by Yagi antenna beside cabin, vertical polarisation, $V = 85\%V_r, f = 60$ Hz, location 1 51
Figure 6.15: PD measurement by Yagi antenna beside cabin, vertical polarisation, $V = 115\%V_r, f = 60$ Hz, location 1 51
Figure 6.16: PD measurement by monopole antenna, $V = 115\%V_r, f = 60$ Hz, location 1 51
Figure 6.17: PD measurement by helical antenna, $V = 115\%V_r, f = 60$ Hz, location 3 52
Figure 6.18: PD measurement by Yagi antenna, location 2, $V = 100\%V_r, f = 50$ Hz Observation throughout 50 voltage cycles 52
Figure 6.19: PD measurement by Yagi antenna, location 1, $V = 85\%V_r, f = 60$ Hz Observation throughout 50 voltage cycles 53
Figure 6.20: PD measurement by Yagi antenna, location 1, $V = 115\%V_r, f = 60$ Hz Observation throughout 50 voltage cycles 53
Figure 6.21: PD measurement by monopole antenna, location 1,V = 115%Vr, f = 60 Hz Observation throughout 50 voltage cycles 53
Figure 6.22: PD measurement by helical antenna, location 3, V = 115%Vr, f = 60 Hz Observation throughout 50 voltage cycles 54
Figure 7.1: Ground plan of tested transformer and arrangement of antennas 56
Figure 7.2: One measurement of 100 pulses captured on the transformer 56
Figure 7.3: The background noise of the signal and its frequency content measured by Yagi aerial 57
Figure 7.4: The background noise of the signal and its frequency content measured by helical antenna 57
Figure 7.5: The background noise of the signal and its frequency content measured by the KEMA probe 58
Figure 7.6: 1st group of signals measured by Yagi 1 and the KEMA probe; triggered on Yagi signals 59
Figure 7.8: 3rd group of signals measured by Yagi 1 and the KEMA probe; triggered on Yagi signals 60
Figure 7.9: Distribution of the number of pulses regarding the voltage phase angle for signals measured by Yagi 1 60

Figure 7.10: Distribution of the number of pulses regarding the voltage phase angle for the 1st (left graph) and the 2nd (right graph) group of signals 61
Figure 7.11: Distribution of the number of pulses regarding the voltage phase angle for the 3rd group of signals; left side – positive cycle and right side –negative cycle 61
Figure 7.12: 1st group of signals measured by the helical antenna H1 and the KEMA probe; triggered on helical antenna signals 62
Figure 7.13: 2nd group of signals measured by the helical antenna H1 and the KEMA probe; triggered on helical antenna signals 62
Figure 7.14: 3rd group of signals measured by the helical antenna H1 and the KEMA probe; triggered on helical antenna signals 63
Figure 7.15: 4th group of signals measured by the helical antenna H1 and the KEMA probe; triggered on helical antenna signals 63
Figure 7.16: Distribution of the number of pulses regarding the voltage phase angle for H1 64
Figure 7.17: Distribution of the number of pulses regarding the voltage phase angle for the 1st and the 2nd group of signals 64
Figure 7.18: Distribution of the number of pulses regarding the voltage phase angle for the 3rd and the 4th group of signals 64
Figure 8.1: Data mining process 68
Figure 8.2: An example illustrating PCA 78
Figure 8.3: Two-dimensional SOM with neurons arranged in a hexagonal lattice 80
Figure 8.4: The neighbourhood relation of hexagonal and rectangular grid 81
Figure 8.5: Influence of input vector towards its BMU and neighbours 82
Figure 8.6: Neighbourhood functions of SOM (left rectangular and right Gaussian function) .. 83
Figure 8.7: Selection of optimum SOMs based on several suitable candidates 85
Figure 8.8: The input data where input vectors consists of three components 86
Figure 8.9: Arrangement of neurons 86
Figure 8.10: The u-matrix and component-planes 87
Figure 8.11: Input data "transformed" into u-matrix 87
Figure 8.12: Distinction between agglomerative and divisive techniques 92
Figure 8.13: Agglomerative method algorithm 94
Figure 8.14: Contraction effect 95
Figure 8.15: Dilatation effect 96
Figure 8.16: Reversal effect 96
Figure 8.17: The input data where input vectors consists of three components 98
Figure 8.18: Dendrograms based on the Euclidian distance and the Ward (left) and the Average (right) linkage method 98

Figure 8.19: The distance between two successive clusters for the Ward (left) and the Average (right) linkage method 98
Figure 8.20: Final clusters obtained by Ward (left) and average (right) linkage method 99
Figure 9.1: Various stages of the proposed approach 101
Figure 9.2: The R measure based on *i* principal components 102
Figure 9.3: The scree plot 102
Figure 9.4: U-matrix and component planes of the SOM 104
Figure 9.5: Chosen clusters assigned in the u-matrix 105
Figure 9.6: Group 1 106
Figure 9.7: Group 2 107
Figure 9.8: Group 3 108
Figure 9.9: Group 4 109
Figure 9.10: Group 5 110
Figure 9.11: Group 6 111
Figure 9.12: Group 7 112
Figure 9.13: Group 8 113
Figure 9.14: Group 9 114
Figure 9.15: Group 10 115
Figure 9.16: Group 11 116
Figure 10.1: Various stages of obtaining groups by CA 119
Figure 10.2: Mahalanobis distance and Euclidian distance 120
Figure 10.3: Clustering results for Euclid similarity measure and Ward linkage algorithm 121
Figure 10.4: Distance between two successive clusters for Euclid similarity measure and Ward linkage algorithm 121
Figure 10.5: PD measurements grouped into five clusters considered Euclid similarity measure and Ward linkage algorithm 122
Figure 10.6: PD measurements grouped into ten clusters considered Euclid similarity measure and Ward linkage algorithm 122
Figure 10.7: Silhouette plot of PD data of 10 clusters for case 1 124
Figure 10.8: Clustering results for Euclid similarity measure and Average linkage algorithm 125
Figure 10.9: Distance between two successive clusters for Euclid similarity measure and Average linkage algorithm 126
Figure 10.10: Clustering results for Mahalanobis similarity measure and Average linkage algorithm 126
Figure 10.11: Distance between two successive clusters for Mahalanobis similarity measure and Average linkage algorithm 127

Figure 10.12: Clustering results for Mahalanobis similarity measure and Ward linkage algorithm 127
Figure 10.13: Distance between two successive clusters for Mahalanobis similarity measure and Ward linkage algorithm 128
Figure 10.14: Silhouette plot of PD data for case 2 128
Figure 10.15: Silhouette plot of PD data for case 3 129
Figure 10.16: Silhouette plot of PD data for case 4 129
Figure 11.1: Input data transfer into groups obtained by SOM 132
Figure 11.2: Input data transfer into groups obtained by Mahalanobis distance measure and Ward linkage algorithm 132
Figure 11.3: Input data transfer into groups obtained by Euclidian distance measure and Ward linkage algorithm 132
Figure 11.4: Average magnitude of frequency spectra of PD signals for 100 MHz windows of group 4 134
Figure 11.5: Detail consideration of group 4 obtained by SOM 135
Figure 11.6: Captured signal and its frequency spectrum belonging to blue circles in Figure 11.5 135
Figure 11.7: Captured signal and its frequency spectrum belonging to red circles in Figure 11.5 135
Figure 11.8: Captured signal and its frequency spectrum belonging to green circles in Figure 11.5 136
Figure 11.9: Comparison of the group of signals indicating oil discharges smaller magnitude obtained by SOM, CA (Euclidian – Ward) and CA (Mahalanobis –Ward), respectively 136
Figure 11.10: Clustering results for Euclidian similarity measure and Ward linkage algorithm, when first three PC are considered 137
Figure 11.11: Distance between two successive clusters for Euclidian similarity measure and Ward linkage algorithm, when first three PC are considered 138
Figure 11.12: 3-dimensional presentation of PD data transferred into PC space when first three PC are considered. Groups are formed for Euclidian similarity measure and Ward linkage algorithm 138
Figure 11.13: 2-dimensional presentation of PD data transferred into PC space when first three PC are considered. Groups are formed for Euclidian similarity measure and Ward linkage algorithm 139
Figure 11.14: Clustering results for Mahalanobis similarity measure and Ward linkage algorithm, when first three PC are considered 139

Figure 11.15: Distance between two successive clusters for Mahalanobis similarity measure and Ward linkage algorithm, when first three PC are considered 140
Figure 11.16: 3-dimensional presentation of PD data transferred into PC space when first three PC are considered. Groups are formed for Mahalanobis similarity measure and Ward linkage algorithm 140
Figure 11.17: 2-dimensional presentation of PD data transferred into PC space when first three PC are considered. Groups are formed for Mahalanobis similarity measure and Ward linkage algorithm 141
Figure 11.18: Comparison of the group of signals indicating oil discharges smaller magnitude obtained by CA (Euclidian – Ward) and CA (Mahalanobis –Ward) and group of signals indicating oil discharges larger magnitude obtained by both methods 141
Figure 11.19: U-matrix obtained by SOM, when only first three CA are used as input variables. 142
Figure 12.1: Oil discharge source 145
Figure 12.2: Floating discharge source 145
Figure 12.3: Transformer winding used for generating discharge sources 146
Figure 12.4: The background noise of the signal and its frequency content 147
Figure 12.5: PD signal caused by floating discharge and its frequency spectrum 148
Figure 12.6: PD waveforms measured at different locations and their frequency spectra 148
Figure 12.7: Single pulse caused by oil discharge and its frequency spectrum 148
Figure 12.8: Multiple pulses caused by oil discharge and frequency spectrum 149
Figure 12.9: The shape of air discharge and its frequency spectrum 149
Figure 13.1: Time-frequency plane of signals captured at the Northfleet substation by KEMA probe (left) and helical antennas (right) 151
Figure 13.2: Time-frequency presentation of signals captured by helical antennas placed at position H1, H2, H3 and H4 around the transformer 152
Figure 13.3: Time-frequency presentation of signals captured at Northfleet and Melksham substations 152
Figure 13.4: Time-frequency presentation of signals captured by the antenna in the Laboratory 153

Index to Tables

Table 3.1: Diagnostic methods 13
Table 3.2: Failure distribution for substation transformers with OLTCs [37] 22
Table 3.3: Measuring quantities and number of sensors [36] 25
Table 3.4: Transformer monitors and its failure modes 26
Table 6.1: Antennas and their frequency range used for PD measurements 47
Table 8.1: Clustering algorithms 93
Table 8.2: Agglomerative methods and their properties 95
Table 9.1: Configuration, training and visualisation parameters of SOM 103
Table 10.1: Four cases with different similarity measures and linkage algorithms. 120
Table 10.2: Number and the percentage of total amount of PD measurements in each cluster and cluster average silhouette width 124
Table 11.1: Agreement of groups obtained by SOM and CA (Mahalanobis-Ward) 133
Table 11.2: Agreement of groups obtained by SOM and CA (Euclid-Ward) 133
Table 11.3: Agreement of groups obtained by Mahalanobis-Ward CA and Euclid-Ward CA. 133

1. INTRODUCTION

1.1. OVERVIEW

Between 1884 and 1885 three engineers at the Ganz factory, Bláthy, Déri and Zipernowsky developed a new current distribution system based on the use of the induction apparatus called the "transformer" [1]. The transformer is a relatively simple device, yet it is this electrical device that has made possible the transmission of power over long distances and its' distribution to industry and to our homes at suitable voltages. Since its invention, it has continued to play an important role in electric power system networks. Generally, power transformers are very reliable elements, but when problems occur they are often difficult to diagnose and expensive to correct. Forced outage of a large transformer due to insulation breakdown may lead to catastrophic failure for the transformer, associated equipment and cause environmental damage. Accurate assessment of the transformer insulation condition is therefore essential for safe and economic operation.

The challenges faced by electric utilities over the past years are unrelenting and can be summed up in one sentence: "Reduce operating costs, enhance the availability of the generating and transmission equipment, and improve the supply of power and service to the customer base". Transformers are critical elements of the electrical power system and their reliable and continued performance is the key to profitable generation and transmission. Monitoring systems for power transformers can help to achieve the foregoing goals. They offer the possibility of extending the life span of power transformers, reduce the risk of expensive failures, and provide potential for improving the maintenance strategy.

From a purely statistical point of view, transformers have low failure rates and long utilization times. To derive the expected service or remaining life time from such statistical data often leads to very inaccurate conclusions. The deviation can be as large as +/- ten years [2]. There are several reasons for this discrepancy. The life expectancy of a transformer depends not only on service data, operational and environmental data, but also on design and manufacturing aspects. Another significant influence on the expected life can be attributed to the utility maintenance concept and the practical methods adopted for carrying out repair and maintenance operations on each unit. A direct consequence of the foregoing is that transformers have to be

considered and evaluated on an individual basis. The 'remaining life' determination of a transformer also involves many uncertain assumptions including unknown future events and uncertain definition of the end of life. This also implies that an absolute determination of the remaining life in terms of years is not possible.

The assessment of the consumed and remaining life of a transformer is a very important issue relating to monitoring and diagnostics today. This is of interest for both old and new transformers, especially if they have been overloaded or overstressed. End of life is reached when a transformer is not able to fulfil its required function [3] which is to convert power from one voltage level to another. The lifetime of a transformer is defined by strategic, economic and technical factors. The *strategic factor* relates to the transformer's ability to carry the loads, short circuit currents, network service voltage, overvoltages and other stresses applied to it in service. The *economic factor* includes the cost of losses, and maintenance costs.

Technical factors are more complex than strategic and economic factors and include ageing, overstressing and contamination. Mechanical overstressing may be caused by current stresses such as overloads, short circuits or inrush currents, which impose electromagnetic forces on the winding structure, leading to displacement and possibly dielectric breakdown. It may also arise due to vibrations caused by transport shocks or resonance phenomena. Electrical overstressing resulting in dielectric breakdown may be caused by atmospheric or switching overvoltages and winding resonances. Dielectric breakdown may also be a secondary effect of mechanical overstressing, ageing or contamination. The technical end of life of a transformer is a function of design, historical events, previous operating conditions, its present status, and of the future service conditions.

It is easily recognized that not only load and temperature, but also the number of short circuits and overvoltages, design weaknesses, repairs, transportation, etc., will influence transformer ability to fulfil its function.

Degradation of power transformer insulation, which mainly consists of oil and paper, is the main factor for power transformer failures. Chemical analyses and electrical measurements are used for monitoring the condition of transformer insulation. Among these, chemical analyses provide direct information on parameters, such as water content, degree of polymerisation of paper, sludge content in the oil, acidity of the oil and quantity of different gasses dissolved in the oil. However, most chemical analyses must be performed under laboratory conditions and for some chemical analyses (e.g. Chromatography tests) paper samples are needed. On the other hand, electrical measurements are simpler and it is possible to perform them on-site. Because of this simplicity, electrical tests are currently preferred for condition monitoring of transformer insulation although they do not provide direct information about the afore-mentioned parameters.

Detection of partial discharges (PD) from onsite or experimental measurement is essential to understand the characteristics of the PD not only emanating from power transformers but also from other power apparatus, such as, gas insulated substations [4], motors [5], generators [6], cables [7], etc. Nowadays PD measurement techniques employed on oil filled power transformers can be subdivided in chemical, acoustical [8]- [11], electrical [12], [13], and optical [14] measurements. Chemical methods are based on the analysis of dissolved gas generated inside the transformer due to PD activity. The integral characteristic of these regularly performed analyses allow indications on the long term behaviour of the PD activity and, therefore, on the insulation condition. For information on the actual PD occurrence, acoustic and electrical PD measurements are preferable. The focus of acoustic or ultrasonic measurements is based on a PD location, whereas the electrical measurements are orientated to a precise determination of the apparent charge. The optical PD measurement is based on the principle of the detection of light caused by ionisation, which is produced by PDs and recombination process during a discharge. The basis for the optical detection is the electromagnetic spectrum resulting from PDs.

In a power transformer, the presence of PDs is often an indication of the degradation of its insulation. PDs can ultimately lead to the electric breakdown, which can lead to a catastrophic failure not only of the transformer but also of the power system. PD occurrence is characterized by increases of the local electric field strength in a range of inhomogeneities. PD can be described as an electric pulse or discharge in gas-filled void or on a dielectric surface of a solid or liquid insulation system. The pulse only partly bridges the gap between phase insulation to ground, or phase to phase insulation. PDs can occur in gaseous, liquid and solid insulations, whereby no short time disturbance to the electric resistance (insulating property) is apparent. Effects on the surrounding dielectric medium are apparent despite the small transformation of energy of single discharges due to their repeated occurrence. Particularly in organic insulations, the electric erosion by PD leads to a significant reduction in the life span. In the view of the fact that PDs have a number of distinct characteristics, this can be used to advantage for their detection and localisation.

The field of PD diagnostic is complex and to date there is still not a single and well-defined test and diagnostic procedure. In a general sense, this topic can be divided into three parts: the first one relates to measurements, the second one relates to recognition and understanding of the processes involved and, finally the last one is associated with the diagnostic of the condition of the transformer. Hitherto, a significant amount of work has resulted in an improvement of measurement instruments. However, with regard to source recognition and identification of specific defects affecting a power transformer, the effort has been much less coordinated and is still the case that different experts taking measurements in the same transformer can come up

with different diagnostics. This, in part, contributes to casting some doubt about the validity of existing PD measurement and diagnostic techniques.

1.2. OBJECTIVE OF THE PROJECT

This thesis focuses on investigation of a non-invasive partial discharge measurement system that can be used for monitoring of power transformers. The principal aims are:

- To investigate the feasibility of transformer PD measurement by a non-invasive measuring system consisting of wideband antenna and oscilloscope.
- To determine, separate and characterise different discharges from measured signals.
- To investigate classical and artificial intelligence methodologies for the purposes of PD classification.

1.3. SCOPE OF THE THESIS

The structure of this thesis is as follows:

Chapter 2 briefly describes transformer construction and its insulation system. The sources and effects of transformer insulation degradation are also reviewed.

Chapter 3 contains an overview of high voltage power transformer diagnostic techniques. In addition to a brief description of existing methods, the effectiveness of various detection and diagnostic methods are presented. Transformer monitoring equipment is also introduced. Typical waveforms of different observable quantities captured via the monitoring equipment are illustrated.

Chapter 4 introduces different types of PDs and their causes. The physical background of PDs is briefly described.

Chapter 5 presents the use of PD measurements as a transformer diagnostic tool. Different PD detection methods and PD pattern analysis techniques are explained.

Chapter 6 deals with the measurement set up and results obtained from the first field trial, based on non-invasive PD measurements. Measurements were performed at the Nielson substation on 400/275 kV power transformer. This chapter concludes with problems encountered within the measuring system employed and gives suggestions for improvement for future tests.

Chapter 7 contains a test set up, PD acquisition system and results of PD wide band measurements performed at the Northfleet substation on the SGT3A 400/275 kV power transformer.

Chapter 8 related to feature extraction and grouping of signals, proposes and examines the use of different data mining (DM) techniques. Specially, this chapter focuses on a principal component analysis (PCA) and the method of moments, as a tool for feature extraction. Additional, self organizing maps (SOM) and cluster analysis (CA) are presented. Advantages and disadvantages of proposed methods are discussed.

Chapter 9 presents the outcome the SOM conducted on the PD data obtained at the Northfleet substation. The applied method enables good formation of PD groups. Finally, the advantages of the proposed approach are presented and discussed.

Chapter 10 deals with the CA implemented on the PD data obtained at the Northfleet substation. Two distance measures and two cluster algorithms are applied and their advantages and disadvantages are discussed. Applied method shows a good potential for grouping PD signals.

Chapter 11 compares results obtained by SOM and CA. Discussion is confused on the advantages and disadvantages of applied methods.

Chapter 12 presents PD measurements performed at UMIST Laboratory. Firstly, three models of PDs defects are illustrated. Secondly, the outcome of the analysis of PD measurements is presented. Finally, the advantages of the proposed approach are presented and discussed.

Chapter 13 explores the potential of another method of DM, the method of moments, applied to PD measurements. Based on the investigations, it is shown, that this method is not very efficacious for PD data considered in this work.

Conclusions and Recommendations are presented in **Chapter 14.** Particular specific conclusions pertinent to the field of transformer PD measurements are summarized. Recommendations are made for future work in the field of non-invasive transformer PD measurements.

2. COMPONENTS AND PROPERTIES OF A TRANSFORMER INSULATION SYSTEM

2.1. TRANSFORMER CONSTRUCTION

A transformer is an electrical device by which alternating current of one voltage is changed to another. The simplest transformer consists of an iron core and two windings inductively coupled with the core. The windings of a two-winding transformer are distinguished either by the service condition (the supply winding is called the primary, and the load winding the secondary) or by their voltage levels (high-voltage winding and low voltage winding).

The core is the part of the transformer in which the electromagnetic field oscillates. It consists of laminated steel. The core concentrates and focuses the flux in the transformer. Copper conductors of transformer winding are insulated with Kraft paper, though the uses of wires with enamel insulations are finding more widespread usage. The paper strip is applied as overlapped over the conductor. The interturn insulation in a transformer winding serves to separate from each other the turns energized to different potentials. The winding insulation or inter-section insulation is generally provided between sections of the disc-type windings of transformer to ensure the required dissipation of heat. The main insulation separates the windings of different voltage ratings and the windings from earthed parts. Between cylindrical windings mounted on a common limb but having different voltage ratings, is the layer insulation which generally consists of oil and some kind of solid insulation. The oil layer is divided into ducts by paper-based phenol tubes or tubes made of pressboard. Insulation of winding ends from each other and from the yoke is the task of the so-called end-insulation.

Today's transformers are almost entirely oil filled; therefore the tank contributes one of the main parts of a transformer. The tank encloses the assembled core and windings. It acts as a protective barrier for the high voltage and current carrying components holds the transformer oil used for cooling and helps dissipate the heat built-up within the transformer.

2.2. TRANSFORMER INSULATION

It is hardly necessary to emphasise the importance of a reliable insulation system for the modern power transformers. Internal insulation failures are invariably the most serious and costly of transformer problems. High short-circuit current levels on today's electrical networks ensure that the breakdown of transformer insulation will almost always result in a major damage to the transformer. However, consequential losses (due to damage to transformer) such as the non-availability of a large generating unit can often be far more costly and wide reaching than the damage to the transformer itself.

This overview is focused on the insulation system of oil impregnated power transformers, in which Kraft paper is used as conductor insulation whereas pressboard and duct filled with mineral oil are mainly used as the main insulation.

2.2.1. Solid insulation

Solid insulations are used in four different parts of the internal structure, viz., between turns, between coils and to separate one winding from another winding, and from all windings to ground.

Kraft paper is a key component of transformer insulation. It is a low cost material obtained from wood pulp with outstanding mechanical and electrical properties. Paper is composed of 90 % cellulose, 6 – 7 % hemi-cellulose and 3 – 4 % lignin [15]. Long cellulose fibres provide high mechanical strength to paper. However, to obtain higher density, which gives higher electric impulse strength and higher dielectric permittivity, the length of fibres should be shorter. Therefore, to achieve good electrical properties, a compromise is needed at the expense of weakening mechanical strength.

The important dielectric properties of paper, which is known for its insulation quality, are dielectric permittivity ε, loss factor tan δ and conductivity σ. The permittivity of dry paper varies from 1.5 to 3.5 and the loss factor varies between 0.003 and 0.004. Dry paper has a very high volume resistivity, i.e. values between 1015 and 1017 Ωcm can be obtained [15]. However, paper has to be protected from direct contact with moisture to maintain its good dielectric properties, due to the high affinity of paper to water.

There are areas in transformers where the electrical and mechanical stresses are high and these cannot be withstood by single layers of paper. Therefore, pressboard is used instead. Pressboard is produced by wet pressing several paper layers without any bonding material. The density of pressboard can reach up to 1.3 g/cm^3 due to its low porosity, compared with that of

paper. Also the permittivity of pressboard is higher than the permittivity of paper, i.e. around 4.5 [16].

2.2.2. Mineral oil

In the majority of power transformers, mineral oil is used as the insulating liquid due to its wide availability and its low cost compared to other equivalent products, like silicon oil and organic esters. The main purpose of using mineral oil is to impregnate the paper insulation to avoid direct contact between paper, air and moisture. The oil also acts as a heat transfer medium, which controls unnecessary rise in the temperature of the apparatus due to losses in conductors, core and dielectric materials [17], [18].

Transformer mineral oil is refined from hydrocarbons collected during the distillation of crude petroleum. Characteristics of the refined oil depend on the relative proportions of paraffinic, napthenic and aromatic hydrocarbons. Of these, the aromatic component is chemically unstable compared with the other two, due to its unsaturated chemical structure. Therefore, aromatic components in mineral oil have to be minimised in the final product using appropriate refining methods. Depending on the refining method, the proportions of paraffin, naphthene and aromatics can vary between 40 – 60 %, 30 – 50 % and 5 – 20 %, respectively [18], [19]. All these hydrocarbons have little or no polarity. Apart from the above named constituents, one can find traces of polar compounds (sulphur, oxygen and nitrogen) and ionic species (organic salts) present in mineral oil. These constituents have an immense influence on the chemical and electrical properties of the mineral oil.

To obtain better performance as a cooling medium and as electrical insulation, transformer oil must have high dielectric strength, low viscosity, high heat capacity and a low expansion coefficient. It should also be free from moisture, gases, chemical impurities and mechanical contaminants to avoid unnecessary electrical discharges.

2.2.3. Oil impregnated paper insulation

Paper itself shows modest dielectric performance due to its porous structure. When air is present, the dielectric strength of paper is predominantly determined by the gaseous ionisation within the air space. Heat, water and oxygen accelerate the degradation of paper insulation [15]. To control the influence of the above mentioned critical factors, it is necessary to impregnate paper insulation.

The impregnation of paper insulation is a sophisticated process. However, the technology does not allow achieving 100 % perfect impregnation. Imperfect impregnation may cause

damage to the insulation during service. This may happen due to the PDs through the gas-trapped voids between paper layers or paper fibres.

Most of the electrical properties of impregnated paper, except for the breakdown voltage, are closely dependent on the electrical properties of dry paper, since the proportionality of oil is small compared with paper and these two materials are chemically inert with each other.

2.3. INSULATION DEGRADATION AND AGEING

2.3.1. Degradation of oil-paper insulation systems

Changes in electrical, mechanical and chemical properties of the insulation due to degradation phenomena during service can lead to severe damage. The rate of degradation of oil-paper insulation depends on the existing thermal, oxidative, hydrolytic, electrical and mechanical conditions within the transformer [15].

The weakest link of any transformer insulation system is the cellulose. Cellulose cannot be restored to new without a complete repair, unlike the oil, which can be replaced as needed. The Achilles heel of cellulose comes from its affinity for water and the compounds of oil degradation. Cellulose is vulnerable to oxygen and to excessive heat. When these elements are present at the same time, the ageing process is accelerated. For example, a 1 per cent water content will age the cellulose 10 times faster than if the water content was 0.1 per cent. If we combine the effect of oxygen in the presence of water, we can obtain an ageing effect in the order of 2.5 times [15]. In the majority of cases, the failure of a transformer is related to an insulation breakdown, and the reason for the failure is usually mechanical; under the stress of physical forces, the insulation gives way. The adjacent surfaces of the spires have to support high mechanical stress. This physical stress comes principally from continuous vibrations and short-circuits. Under the effect of a short-circuit, windings undergo enormous pressure. These forces cause vibrations in all parts of the windings. If the cellulose is in good condition, the insulation will support the effect of mechanical stress, and if not, the latter will be damaged or destroyed, causing a latent default or a transformer failure.

Not only the paper but also the properties of oil are affected by the presence of oxygen. Oxidation of napthenic and parafinic hydrocarbons increases the acid number, whereas oxidation of aromatics remarkably increases the loss factor. The rate of oxidation of oil depends highly on temperature, the presence of light and catalysts. At elevated temperatures and with the presence of active metals, such as Cu, Pb and their alloys, rapid oxidation of oil can be experienced.

Ionisation is the other important process that increases the ageing of oil-paper insulation. During initial processing of oil-paper insulation, it is not possible to remove whole trapped gases within the insulation. However, during service, these local areas, where air bubbles are trapped, may be stressed by high electric fields. The field strength in the bubble can be higher than the breakdown strength of air and can lead to discharges within the insulation. Oil molecules absorb energy from these discharges and decompose into hydrocarbons and hydrogen. If oil is saturated with gases, this hydrogen may create more bubbles, which develop more discharging paths within the insulation. Therefore, the breakdown strength of insulation decreases with time. Also discharges initiated in cavities of paper insulation produce conductive carbon particles [18].

3. ASSESSMENT OF POWER TRANSFORMER CONDITION

3.1. INTRODUCTION

The diagnostics of a transformer contains sophisticated off-line measurements together with data interpretation. The diagnostics is normally used either for determining the actual condition of a transformer i.e., its level of degradation, or as a response to a received warning signal.

There is a distinct difference between the preventive diagnostics and fault diagnostics; the former determines the time at which different parts of transformer are satisfying required standards, and the latter determines the extent of a fault and its cause.

The diagnostics includes different measurement techniques, procedures, as well as expert knowledge. It is a consequence of theoretical, experimental and empirical knowledge in relation to the power transformer operation phenomena. Of course, these methods are dependent on each other.

Some of the diagnostics measurements are performed on-line and the other off-line. On-line measurements are carried out during regular operation while off-line methods require disconnection of a transformer from the power network and are mainly used during scheduled inspections or when the transformer is already suspected to be defective. A brief survey of electrical chemical and physical methods and their diagnostic results is given in Table 3.1. The operating condition of a transformer and the fault location capability are also included. This chapter briefly reviews several techniques used in the condition monitoring of power systems.

3.2. ELECTRICAL METHODS

3.2.1. Leakage inductance measurement

This method is used for the detection of mechanical deformation of windings, caused by strong electro-mechanical forces during short-circuit. The mechanical forces affect the windings and deform them. A change in inductance is a consequence of geometrical change of the winding. The leakage inductance change can be caused by turn-to-turn faults or by winding

deformation. In order to obtain information about the winding condition, previous results must be compared with measurements obtained. If the leakage inductance changes between 0.5 to 2 %, action needs to be taken, because that could be a possible reason for a winding deformation [20]. The change of more than 2% could mean an alarm. If there is any doubt about the obtained results, one should also consider the results obtained from measurement in time or frequency domain.

3.2.2. Measurement of magnetizing currents

The magnetizing currents are measured at low voltage (usually 380 V) in all the three phases. Reliability of this method is only good for major core and winding deformations and for checking the connection group of a transformer. In all other cases, its detection capability is very low. However, the advantage of this method is in its simplicity and low price. The results from other methods such as dissolved gas analysis and leakage inductance measurement have to be considered to complement this method.

3.2.3. Insulation resistance testing and polarization index

Insulation resistance (IR) measurement is one of the conventional methods used for determining the dryness of the transformer insulation. IR is measured by applying a fixed DC voltage, usually 0.25-5 kV, across the insulation. The resultant current, which is a combination of capacitive current, absorption current and conduction current, monotonically decreases. When a DC voltage is suddenly applied to insulation, the insulation current will start at a high value, gradually decrease with time, and finally level off to a stable value. The low initial insulation resistance is partly caused by the high initial capacitive charging current. This capacitive current rapidly decreases to a negligible value (usually within 15 seconds) as the insulation becomes charged. The low initial insulation resistance is also partly caused by the high initial dielectric absorption current. This current also decreases with time, but more gradually, requiring from 10 minutes to several hours to decay to a negligible value. However, the change in dielectric absorption current after 10 minutes can be disregarded. The leakage current does not change with time when a voltage is applied, and this current is the primary factor on which insulation quality may be judged. Insulation resistance varies directly as the thickness and inversely as the area of the insulation being tested. A curve plotted between insulation current and time (or insulation resistance and time) is known as a dielectric absorption curve.

Table 3.1: Diagnostic methods

Diagnostics method (test)	Operation condition	Location of fault	Diagnostic result
Electrical methods			
Leakage inductance measurement	Out of service	Partial (can indicate the faulted phase and the faulted winding)	Mechanical deformation of the windings (detecting changes in the winding geometry)
Measurement of magnetizing currents	Out of service	Partial (can indicate the faulted phase and the faulted winding)	Rough deformation of Fe core and winding faults (Shorted turns and some core faults)
Recovery voltage measurements (polarisation spectrum)	Out of service	None (overall measurement), although may indicate presence of wet spots	The humidity content of the isolated material (paper moisture/ageing)
Testing of insulation resistance and polarization index	Out of service	Partial (can indicate the faulted phase and the faulted winding)	The content of moisture in isolation and detection of weak places in insulation
Dissipation (or loss) factor measurement and capacitance (Power factor)	Out of service	Partial (can indicate the faulted phase and the faulted winding)	An overall assessment of the insulation
Frequency response analysis	Out of service	Good (can identify the faulted phase and the faulted winding)	Mechanical deformation of the windings
Partial discharge measurement	Out of service	Partial	Dielectric breakdown
Chemical methods			
Chemical analysis of the oil (IEC 422)	In service	None	An overall assessment of the status of the insulating oil (dielectric strength, moisture content, loss angle, resistivity, containment of inhibitor)
Dissolved gas analysis	In service	None (integrates over time)	Thermal and dielectric stress of oil and paper (discharge, overheating)
Degree of polymerisation	Out of service	None	Quantify the degradation of paper insulation (cellulose)
High performance liquid chromatography (Furfural analysis)	In service	None (integrates over time)	Quantify the general degradation of the cellulose insulation (furanic compounds analysis)
Physical methods			
Radio interference	In service	None or partial	Discharge
Acoustic emission	In service	Good if discharge is not in winding	Discharge
Vibration	In service	Partial	Core and winding deformation
Infrared emission	In service	Good	Tank currents, Cooler blockages

The polarization index gives information about the moisture content of insulation and is defined as follows:

$$I_{P1} = \frac{i_{15"}}{i_{60"}} = \frac{R_{60"}}{R_{15"}}$$
$$I_{P1} = \frac{i_{1'}}{i_{10'}} = \frac{R_{10'}}{R_{1'}} \tag{3.1}$$

where $i_{15"}$ is current measured after 15 seconds, $i_{60"}$ and $i_{1'}$ are currents after 60 seconds, and $i_{10'}$ is current measured after 10 minutes.

The main parameter obtained with this method is the insulation resistance $R_{10'}$, which is determined as:

$$R_{10'} = \frac{U_0}{i_{10'}} \tag{3.2}$$

The higher is the valued $R_{10'}$, the better is the condition of the dielectric. Typical good values of polarisation index may be 1.5-2.0 (poor values 1-1.2) and can be as high as 5.0 [20]. The polarisation index is used as a measure of dielectric deterioration with time. The advantage of this method is that it also detects weak spots in insulation compared with loss angle measurement.

3.2.4. Recovery (return) voltage measurement - RVM

The RVM is one of the time domain techniques for studying slow polarisation processes in insulating material [19], [21]. It determines the water content of the oil impregnated paper insulation and the absorption of by-products of ageing.

In the first instance, applying a step DC voltage for a specified time period charges the object. During this period, the polarisation current flows through the test object. Subsequently, the test object is short–circuited (grounded) for a certain time (usually shorter than the charging time) and the depolarisation current flows. When the short–circuited period is completed, the recovery voltage $U_R(t)$ is measured under open-circuit conditions. The source of recovery voltage is relaxation of the remaining polarisation in the insulation system, giving rise to an induced charge on the electrodes. The so-called polarisation spectrum is established by performing a series of recovery voltage measurements with stepwise increasing charging time t_c, and short-circuit time t_g, usually with the ratio $t_c/t_g=2$. For each sequence, the peak of the recovery voltage U_{Rmax} and the initial rate of rise of the recovery voltage dU_R/dt are recorded and plotted against the charging time used.

The maximum value of the return voltage is directly proportional to the polarisation ability of the dielectric material, and the initial slope is proportional to the polarisation conductivity, i.e. a higher initial slope of the response as the material ages and degrades. A new transformer

with very dry insulation will exhibit a dominant time constant of several hundred or thousands of seconds, whereas a transformer with fairly high moisture content will have a dominant time constant of the order of seconds. The RVM method effectively completes the range of conventional methods, e.g. loss-factor, dissolved gas analysis, etc.

3.2.5. Polarisation and depolarisation currents (PDC)

Polarisation and depolarisation currents measured between low voltage (LV) and high voltage (HV) windings of oil-paper insulated power transformers have been used for assessing the condition of the insulation [19]. For measuring the polarisation current, fixed DC voltage (over 500 V) is applied across the separately short-circuited LV and HV terminals. The depolarisation current is then measured by short-circuiting the HV and LV terminals through an electrometer after DC voltage has been removed. A guard terminal is also exploited for avoiding the influence of leakage currents.

The major advantage of this method is that more information can be gathered compared with the traditional IR and PI measurements. In-between two consecutive DC measurements, the end terminals have to be short-circuited for a sufficiently long time to reduce memory effects. A simple rule of thumb on time domain measurements is that the short-circuit should last for at least as long as the previous charging time before starting a new measurement. Very sophisticated noise suppression methods are needed for PDC measurements, which is another drawback to using this method as an on-site diagnostic test.

3.2.6. Loss-factor (dissipation factor, tan δ) and capacitance measurement

The loss-factor of insulation is a measure of its dielectric power loss, and is not a measure of its dielectric voltage strength. This method yields a result of overall assessment of the insulation i.e., the degree of degradation. The loss-factor of insulation is defined as the ratio between resistive and capacitive currents caused by an AC voltage applied across the insulation. Subsequently, the total loss of insulating material is characterised by the loss tangent.

In conventional loss factor measurement, the Schering bridge [22] coupled with a high voltage AC source is used. Test voltage used in typical field test sets varies broadly (10 V-12 kV).

An increase in the dielectric loss may accelerate insulation deterioration because of the increased heating, but more commonly, an increase of dielectric loss is evidence of other deterioration, which also affects dielectric strength. As in insulation resistance tests, the change in periodic test readings is more indicative of insulation deterioration rather than the absolute magnitude of readings.

The loss-factor method is particularly useful for detecting moisture and other loss-producing contaminants in transformer windings. Experience has shown that the loss-factor test is more revealing than the insulation-resistance test when there is a high-loss dielectric in series (as in a transformer winding surrounded by oil), and is less influenced by surface leakage components. Users have found many cases where high loss-factor readings have indicated moisture in the windings, although the oil dielectric tests have been up to standard. The loss-factor was brought back to normal by a dry-out run. Time-saving techniques have been developed whereby the losses in transformer bushings or windings can be segregated without disconnecting the windings from the bushings. All windings not at test potential should be grounded. Each winding should be measured to all other windings and to ground, and also all of the windings measured together to ground. All possible winding combinations should be tried.

3.2.7. Frequency response analysis (FRA)

Transformer winding deformation can be measured via the method of frequency response analysis (FRA). The basis of this method is that a change in the physical condition of the transformer, caused by damage or deterioration, would be reflected in the electrical parameters, these being resistance, inductance and capacitance, when hence can be detected by testing the unit's response to a range of input signals of varying frequency [23].

Once a transformer winding is deformed, the insulation will be affected. The extent of insulation damage depends on the deformation. Even if the insulation is not damaged seriously, a mechanically unstable winding is bound to be damaged by the next short circuit impact. It is possible too that a winding with deformation will cause failure by only normal electro-dynamic forces, such is the importance of transformer windings deformation.

When high frequency signals are imposed between a winding terminal and earth terminal, a transformer can be simplified to a 4-terminal (or 2-port) network made of resistors, inductors and capacitors (RLC). The space dispositions of a winding, including its own structure, the distance between the winding and the shell and core of a transformer, which are earthed, decide the distribution of dispersal inductance and capacitance directly. That means that each transformer has its own frequency response characteristics. Once any part of the inner structure of a winding has been changed, its transfer function would definitely change. The winding deformation can be detected through the analysis and comparison with the frequency response characteristics of the winding.

3.2.8. Partial discharge (PD) measurement

Major sources of PD activity in power transformers are defects caused by the mechanical deformation of windings, deterioration and ageing of the components and defects of the

insulating structure of the tap changer. PD phenomena and PD measurements are described in detailed in Chapters 4 and 5.

3.3. CHEMICAL METHODS

3.3.1. Oil analysis

Chemical and dielectric analyses, which are used for estimation of the operating properties and degradation of oil, include the breakdown voltage, water content, dielectric dissipation factor and resistivity, interfacial tension, oxidation stability, total gas content, flash point, pour point, density, viscosity and inhibitor content (which is employed for inhibited oils only) [24].

Power factor measures the amount of energy dissipated as heat. This test helps to reveal the quality and integrity of the insulating liquid and can be used to determine when a filtering or change of transformer liquid is needed.

Interfacial tension determines the presence of polar compounds, which are considered to indicate oxidation contaminants or deterioration from the transformer construction materials.

Acidity measures the acid levels in the transformer liquid. As acid levels increase, the quality of the oil decreases, causing dielectric loss, corrosion, and thermal problems from sludge.

Colour is an important indicator of quality, ageing and the presence of contaminants.

Specific gravity verifies the weight and type of insulating fluid.

Metals-in-oil testing analyzes the presence of 22 metals to provide vital information about fault locations and excessive wear in the mechanical components.

3.3.2. Dissolved gas analysis (DGA)

Transformer oils perform at least four functions for the transformer. Oil provides insulation, provides cooling, and helps extinguish arcs. Oil also dissolves gases generated by oil degradation, moisture and gas from cellulose insulation, deterioration, and gases and moisture from whatever atmosphere the oil is exposed to. Close observation of dissolved gases in the oil, and other oil properties, provides the most valuable information about transformer health.

The measurement of gases in oil is the most established diagnostic method for transformers and it has been applied to transformers in service for more than 30 years. The method is based on analysing the types, concentration and production rates of generated gases. Analysis of the faulty gas mixture dissolved in insulating oil of a power transformer is an effective method for the detection of incipient faults. The gases due to faults in transformers are generally produced by oil and other insulating materials, e.g. cellulose degradation. The most significant gases generated by the decomposition of oil are hydrogen (H_2), methane (CH_4), ethane (C_2H_6),

ethylene (C_2H_4), and acetylene (C_2H_2). Where cellulose materials are involved at the point of a fault, further gases, chiefly carbon dioxide (CO_2) and carbon monoxide (CO) are also generated.

The dissolved gas may vary with the nature and severity of different faults. By analysing the energy density of faults, it is possible to distinguish three basic fault processes:

- overheating (pyrolysis),
- corona (partial discharge),
- arcing discharge.

Corona and arcing arise from electrical faults, while overheating is a thermal fault. Different gas trends lead to different fault types, where the key gas is identified by different methods as for example:

- pyrolysis or oil overheating: $C_2H_4^*$, H_2^*, CH_4^*, C_2H_6,
- cellulose insulation breakdown: CO^*, CO_2^*, CH_4^*, $C_2H_4^*$, H_2, C_2H_6
- corona or partial discharge: H_2^*, CH_4^*, C2H6, CO
- arcing: $C_2H_2^*$, H_2^*, CH_4, C_2H_4.

Directives for results obtained from DGA are given in IEC standard 60559 [25]. In this publication, the rated values of gas contents for different kind of power transformers and levels at which dissolved gases are significant are also available. For diagnosis purpose, relative concentrations of individual gases dissolved in oil are used quite often. They are known as Dörnenburg's ratios, Roger's ratios, IEC's ratios, CIGRE's method and Duval's triangle and these belong to conventional interpretation scheme [26]. Based on the results obtained from DGA, the diagnosis of insulation condition of a transformer can be ascertained. The diagnosis must of course consider the previous transformer results.

3.3.3. Degree of polymerization (DP)

To assess the quality of the paper insulation within transformers, one makes use of a factor called the degree of polymerisation DP, which represents the average number of glucose rings in the cellulose polymer. DP is determined by measuring the intrinsic viscosity of a paper solution in an appropriate solvent. The influence of temperature and the presence of oxygen cause the molecular chains to be broken and this influences the lower degree of polymerisation. From a physical point of view, the paper becomes brittle and loses its mechanical stability, but from a electrical point of view, this means a lower resistance to the dynamic forces during short circuits. The degree of polymerisation is 1100 to 1200 for fresh Kraft paper. After factory drying tests, this value will drop to 800 to 1000 and then decrease exponentially with time. A

* asterisk denotes the characteristic gases

tentative value of about 150 is used to indicate the end of the useful life of the cellulose [27]. For partial rewinding of failed coils, a minimum value of 250 has been used.

This method is good for a quantitative measurement of the insulation ageing. The test however is intrusive, it requires a sample of the cellulose, and therefore is limited to cases where an internal investigation can be performed. The results of the test are used to determine the future life of a transformer and to predict how similarly operated units are performing. Methods that sample the oil and look for degradation products of the paper, such as furan type compounds, are under development at the present time.

3.3.4. High performance liquid chromatography (HPLC)

Cellulose paper in the oil-paper system produces carbon oxides, sugars and furanic by-products, when it is thermally degraded. Furfural analysis is detected by high performance chromatography. HPLC has gained increasing popularity as a means of estimating degradation of insulating papers in power transformers [28] and [29]. The cellulose fibres in paper degrade to produce a range of furans, which dissolve in the impregnating mineral oils. The advantage of using furan levels, instead of carbon oxide levels is, that the former are degradation products specific to the paper. Oil degradation will not produce furans.

Furan is a heterocyclic aromatic system consisting of four carbons and one oxygen in a five-member ring with each of the carbons having hydrogen attached; hence the molecular formula is C_4H_4O. The concentration of furan compound gives an indication of paper condition, and the rate of change of furan compounds concentration can indicate the rate of ageing of the paper. The advantage of this method is that the information about the paper condition in a transformer is obtained without disconnecting it from the network.

If a good correlation can be found between the average degree of polymerization and the concentration of the furanic compounds individually or in combination, then this would have a number of advantages for evaluating the condition of the cellulose insulation within a unit. First by it is a non-intrusive procedure and second by it does not require interruption of service to obtain a sample, finally the laboratory analysis requires less time and it is a more sensitive determination.

3.4. PHYSICAL METHODS

3.4.1. Radio interference

The radio interference meters are used for PD measurements [30]. They are frequency selective voltmeters and are primarily intended for measuring interference caused to reception

of broadcast radio signals. Because of their special characteristics, radio interference meters give a general indication of discharge magnitude when used on the quasi-peak setting.

3.4.2. Infrared emission

Infrared thermography is a valuable predictive maintenance tool. It is applied to the transformer for the purpose of locating abnormal heating. A thermal scan of a transformer main tank and an associated load tap changer can reveal load tap-changer problems, blocked radiators, hot spots, or poor connections. Thermography will detect and translate heat emissions (infrared waves) into electronic signals that are imaged, measured and then analysed. The inspection is fast, accurate and practical.

3.4.3. Acoustic emission

Infrared thermography will not detect corona, arcing and tracking in their early stages, first because the heat generated is non-existent or minimal, and second because the technology is blind to what is going on behind a sealed cabinet. Infrared is a valuable tool only when corona, arcing and tracking go undetected and the condition exacerbates, causing a galloping increase in temperature.

An ultrasonic instrument (such as UE Systems, Inc.'s Ultraprobe 2000 [31]), on the other hand, picks up the sound of corona, arcing and tracking even above the ambient noise of the manufacturing plant and has the ability to "see" through walls. This makes it an ideal tool in its early detection of potentially destructive electrical disturbances.

An ultrasonic instrument detects the high-frequency noise produced by electrical discharges and translates it, via heterodyning, down into audible ranges. An inspector listens to the specific sound quality of each type of emission over headphones while he observes the intensity of the signal on the instrument's meter.

3.4.4. Vibration

Internal transformer problems that may be in the process of developing because of loosening in the core/coil assembly can be identified by vibration detection. Vibration detection provides a non-intrusive testing technique for assessing equipment condition during operation. This diagnostic is performed with a portable battery powered device that can display a response frequency waveform of the vibration of the transformer core assembly.

3.5. CONCLUSION ON DIAGNOSTIC METHODS

Methods for identifying the condition of a power transformer after a fault in a power system or during preventive maintenance have attracted great research interest. While the most commonly applied method is dissolved gas analysis, other useful methods include liquid chromatography, insulation resistance measurement, recovery voltage measurements, loss factor measurement, transfer function techniques, etc. All of these methods are imprecise and require some experience in order to correctly interpret observations. They also require a comparison of data with previous measurements. The best conclusion regarding power transformer condition is to aggregate information from more than one technique described above.

3.6. POWER TRANSFORMER MONITORING

3.6.1. Introduction

Monitoring equipment, which by definition is on-line and permanently mounted on the transformer, can extend the lifetime of a transformer or prevent its catastrophic failure. Monitoring is defined as on-line collection of data and includes the sensor development, measurement techniques for on-line applications, and data acquisition [32]. There are two main reasons for installing the monitoring equipment on a transformer:

1. By monitoring of important functions of the transformer, developing faults can be detected and eliminated before they lead to its catastrophic failure,
2. Allowing for a change from the periodic to condition based maintenance, i.e. maintenance on request.

A large number of transformer monitors designed to monitor one, two or more specific parameters are already on the market [33]-[41]. The monitoring systems, even if on-line, cannot offer any protection against fast-developing problems, but they serve as a good indicator of incipient faults and offer the possibility to reduce their consequences. Selection of a monitoring device must be based on the failure statistics, and on estimated consequences of failures. Another important factor to consider when choosing a monitor is where the information is received and how it is transmitted.

Table 3.2 shows a typical failure distribution for substation transformers with on load tap changer (OLTC) [37]. With the reference to the Table 3.2, the main parts of the transformer to be monitored are the windings, including the main insulation, and OLTC. The most important parameters related to the windings and the insulation system, are gas-in-oil, temperature and PDs. For the OLTC, monitoring of temperature and load is to be regarded as base information.

Table 3.2: Failure distribution for substation transformers with OLTCs [37]

OLTC	13%
Windings	29%
Tank/fluid	13%
Terminals	29%
Accessories	5%
Core	11%

3.6.2. Gas-in-oil

The gases dissolved in oil due to faults are the first signs of a developing transformer fault. Thermal degradation of oil-impregnated cellulose produces carbon monoxide (CO) and carbon dioxide (CO_2). Hot spots in windings, on insulated leads and in areas using pressboard, cellulose components and spacers also produce CO and CO_2. The degradation of oil from PD produces a high concentration of hydrogen (H_2), while arcing produces both H_2 and acetylene (C_2H_2). Overheating of oil also produces methane (CH_4), ethane (C_2H_6) and ethylene (C_2H_4). The fault temperature influences the gases that are generated by breakdown of transformer oil and thus indicates transformer problems.

An on-line monitor that detects at least some of the listed gases may be a useful tool to alert the maintenance personnel about sudden changes in gas production in oil samples.

Hydran by GE [38], gives an alarm on increasing quantities of dissolved gases H_2, CO, C_2H_2, and C_2H_4. When an alarm is triggered, it is used as an indication to take an oil sample from the transformer and analyse it for actual concentration of dissolved gases (see chapter 3.3.2).

Calisto, made by Morgan Schaffer Systems [39], is a device that continuously measures the concentration of dissolved hydrogen and water in transformer oil.

Serveron [40] offers more complex gas-in-oil device called *TrueGas* that is an on-site transformer dissolved gas analyser. *TrueGas* continuously samples, measures, and displays the concentration of eight transformer fault-indicating gases: C_2H_2, CH_4, C_2H_6, H_2, O_2, C_2H_4, CO in CO_2.

3.6.3. Partial discharges (PDs)

The ultimate failure of a transformer involves breakdown of some part of the electrical systems windings, leads, tap changer or bushings. All such phenomena are voltage-related. An incipient failure is caused when voltage stress is applied across some weak point, which can be due to design, excessive water, particles, damage resulting from a short circuit, deterioration due to the overvoltages, etc. The PD may also result from arcs between the metal parts, such as loose connections or floating metal pieces. Since PD usually occurs before complete

breakdown, PD monitoring provides a warning to remove the transformer from service before catastrophic failure occurs. Because PD can deteriorate rapidly to breakdown, it is helpful to monitor it continuously. The three indicators, chemical, electromagnetic and acoustic, characterize the PD effects.

The presence of H_2 in the transformer oil is a clear indication of PD, but it provides almost no indication of where the fault is located.

Electrical PD monitoring devices involve a minimum of three pulse detectors, located on the capacitor taps of each phase bushings. PD within a transformer is located by triangulation technique. Electrical PD monitoring has been difficult to apply in the field due to difficulties in separating the internal from external PD sources. Electrical PD monitoring is, however, possible by *ICM monitor* (Power Diagnostix Systems GmbH [41]).

The important benefit of acoustic monitoring is its ability to locate PDs. If sensors are located at many points around the tank (up to 50) the signals received from several sensors can be used to locate a fault through a process of triangulation. Physical Acoustic Corporation supplies a complete package of hardware and software, known as *Spartan-AT* [42]-[43], to locate PDs in the field.

3.6.4. Ultrasonic and vibration detection

The ultrasonic detection techniques are based on analysis of sounds that the ear cannot normally detect. Electrical arcing, corona discharge, and pressure vacuum leaks can be pinpointed with the aid of detecting devices.

On-line *Vibro-Acoustic Analysis* (Cuttler Hammer) [44], based on the piezoelectric accelerometer, is magnetically attached to the transformer walls and detects transformer core and winding clamping pressure drop and winding vibrations before they lead to accelerate insulation deterioration in coils and laminations.

3.6.5. Temperature

The hot spot of the winding is the limiting factor for assessing the load capability of the transformer. The conventional temperature measurements are not direct; the hot spot is indirectly estimated from the measurements of oil temperatures and of load current. As an alternative, the fibre optic temperature sensors can be installed in the winding when the transformer is manufactured.

Transformer Advantage [45] by Weschler Instruments is a temperature monitor and cooling system controller. Transformer Temperature Monitor (*TTM*), Barrington Consultants Incorporated [15], monitors both oil and winding temperature. In addition to temperature

monitoring and cooling control functions, the "lookback" feature is provided for ambient compensation.

The Fluoroptic Thermometer *(WTS-11)* from Luxtron [46] is a fibre optic transformer winding temperature monitor for direct measurement of transformer hot spot temperatures. It provides an accurate winding temperature during peak loading or emergency overloading of the transformer, as well as safe, cost-effective temperature monitoring in the HV power transformers. *WTS-11* estimates transformer loading based on true winding temperatures and directly measures the winding temperatures or oil temperatures.

3.6.6. Water in oil or cellulose

The water is in a constant state of flux, moving between the cellulose insulation, the oil and the gas space as the pressure and temperature within the transformer change with load. As the transformer warms up, moisture migrates from the insulation into the oil fairly quickly, but returns to the solid insulation more slowly. A transformer that is heavily loaded during the day, cools down rapidly at night, and will exhibit relatively very high moisture concentration or even free water in the oil. The ageing process creates water molecules in the insulation system in the oil. Therefore, the detection and the assessment of the water content is another method for condition monitoring.

The instruments to measure water content of transformer oil are made by Doble Engineering, *Domino* [47] and Panametrics *Probe* [48]. Device *Calisto* (Morgan Schaffer Systems [39]) measures water in oil in addition to the concentration of dissolved hydrogen.

3.6.7. On load tap changer

The OLTC failures are dominated by faults of a mechanical nature, followed by electrical faults and insulation problems. Monitoring of the motor current of the drive mechanism gives indications of certain types of mechanical faults. The temperature difference measurement between the oil in OLTC compartment and the main tank is another method suited for detecting bad contacts and problems of thermal/dielectric nature. Mechanically related problems are not detectable by this technique alone. Recently, vibration monitoring has been applied as a method for diagnostics of OLTCs state. Mechanical and electrical faults can be detected, as well as wear of contacts and changes in transition times.

One of the applications of transformer OLTC monitors is *LTC-MAP* (Load Tap Changer-Maintenance Action Planner) made by GE Harley [49]. Beside the temperature, it monitors the mechanism of the tap changer, tap position, load and motor current.

A similar system, known as *TDM System 3*, manufactured by Barrington Consultant Inc. [50], is also available. It measures tap changer oil temperature and monitors the temperature difference between the temperature of oil in the main tank and in the tap changer compartments.

3.6.8. Integration

Nowadays, producers offer systems that monitor more variables and this enables the constant monitoring of the complete transformer. Integration system for transformer monitoring such as ALSTOM *MS2000* [51], *Faraday* [38], *T-monitor-MAP 22*30 [49], *T-monitor* [53] and *SITRAM*+ [54], measure quantities given in Table 3.3. It is interesting to note that all the listed systems use the Hydran sensor for gas-in-oil monitoring, while moisture and temperature sensors are from different producers. These systems also contain units for capturing data, communication links, and sometimes even their own server, apart from sensors themselves. Producers also offer the software that is capable of dealing with the transformer condition and transformer life time coupled with user friendly interface. An example of an integrated system is given in Figure 3.1.

Table 3.3: Measuring quantities and number of sensors [36]

Measuring quantity	No. of sensors
Currents	1-3
Voltages	1-3
Tap changer position	1
Temperatures (hot spot, ambient, top oil temperature, bottom oil temperature)	1-9
Gas in oil	1-4
Moisture of the oil	1
Oil level in the tap changer tank	1
Oil level in the compensator	1
Operating condition of pumps and fans	4+4

3.6.9. Summary on the transformer monitoring equipment

If a transformer does not have an effective real-time monitoring of its service condition, failures can occur without warning, resulting in service interruption. However, damage to the equipment will be restricted when incipient faults are detected and timely action is taken. Early detection limits the amount of adjacent damage and confines the area requiring repair and maintenance. The relationships of different monitored elements with failure modes in the transformer are presented in Table 3.4.

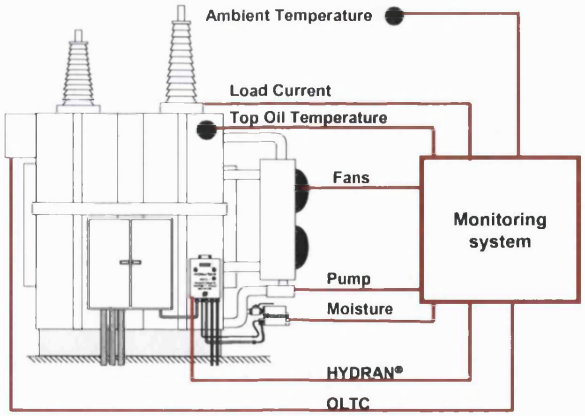


Figure 3.1: Real time inputs to the monitoring system

As can be seen from the table, monitoring of dissolved gases in the transformer covers nearly all failures which can happen in different parts of the power transformer. Observation of PDs helps to detect damage in transformer winding, insulation defects, arcing, gas bubbles and oil contamination.

Table 3.4: Transformer monitors and its failure modes

Observable quantities	**Failure modes and failure indicators**						
	Damage in winding	Defect in insulation	Core damage	Hot spot	Arcing	Gas bubbles	Oil contami-nation
Gas-in-oil	x	x	x	x	x	x	x
Partial discharges	x	x			x	x	x
Thermal- overheating	x		x	x			
Oil breakdown				x		x	x
Moisture-in-oil				x			
Vibrations	x		x				

4. PARTIAL DISCHARGES

4.1. INTRODUCTION

Partial discharges (PD) are small discharges caused by strong and inhomogeneous electrical fields. The reason for such fields can, for example, be voids, bubbles or defects in an insulation material. Detection of PD is performed in order to ascertain the condition of the insulating material in high voltage (HV) elements, e.g. generators, transformers, gas insulated substations (GIS), cables, etc. Since PD usually occurs before complete breakdown, PD monitoring provides a warning to remove the power system element from service before catastrophic failure occurs. In general, there are six types of PD [55]:

1. *Internal discharges* occur in gas-filled cavities, but oil-filled cavities can also break down and cause gaseous discharges afterwards (Figure 4.1).
2. *Surface discharges* may occur in gases or in oil if there is a strong stress component parallel to the dielectric surface Figure 4.2).
3. *Corona discharges* occur at sharp points protruding from electrodes in gases and liquids (Figure 4.3).
4. *Discharges in electrical trees* can start from sharp conducting particles or from a cavity in solid insulation and they may be considered as internal discharges of specific origin (Figure 4.4).
5. *Floating particle discharges* occur when metallic particles "float" in the oil.
6. *Contact noise* may occur in cases of bad contacts or poor grounding of the test samples and tend to concentrate around the zero points of the sine wave, although they are difficult to distinguish from discharge impulses. In transformers, a tap changer or a poorly connected earth foil of a bushing may cause discharges which are not dangerous for the dielectric but nevertheless, are larger than discharges which would be tolerated in the rest of the sample.

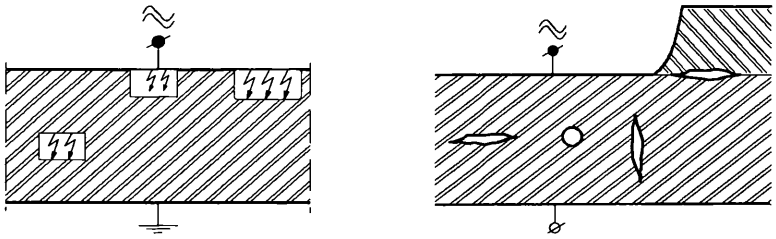


Figure 4.1: Internal discharges

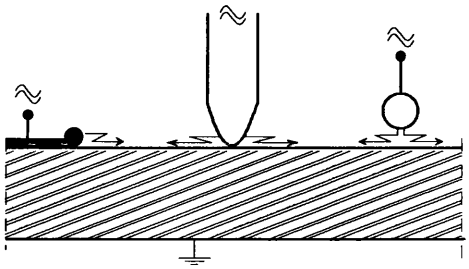


Figure 4.2: Surface discharges

Figure 4.3: Corona discharges

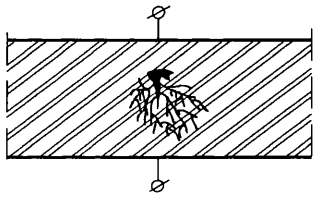


Figure 4.4: Discharges in electrical threes

4.2. OCCURRENCE OF DISCHARGES [56]

4.2.1. Inception of internal discharges

Internal discharges occur in inclusions of low dielectric strength. Usually these are gas-filled cavities, but oil-filled cavities can also break down and cause gaseous discharges afterwards. A frequently occurring inclusion is the gas-filled cavity. It can occur in extruded plastics, cast resins, lapped resin-impregnated paper, etc. The voltage, at which the discharge starts, depends on the stress in the cavity and the breakdown strength of the cavity. In extruded

plastics or cast resin, various inclusions can occur consisting of dirt, paper or textile fibres and other foreign particles. The foreign bodies are more or less impregnated by the plastics, and depending on the kind of plastics used, and inclusions, reduced strength is formed. After breakdown of the inclusion, gas is formed and gas discharges occur. Cavities filled with oil are found between layers and the butt gaps of oil-impregnated paper insulation, such as in transformer windings and cables.

4.2.2. Inception of surface discharges

Surface discharge may occur if there exists a stress component parallel to a dielectric surface. This applies to bushings, end of cables, the overhang of generator windings, and where a discharge from outside touches the surface. The discharge affects the electric field so that, in general, the discharges extend beyond the region where the original surface component of the electric field was high enough to cause discharges.

4.2.3. Inception of corona discharges

Corona discharges occur in gases at sharp points in the electrical field. They occur usually on the high-voltage side, but sharp edges at earth potential, or even at half-way, the electrodes may also cause corona discharges. Although the distance between electrodes may be large, the field concentration at a sharp edge, nevertheless, causes a partial break-down of the surrounding air.

4.2.4. Discharges by electrical treeing

Electrical trees start from defects in the insulation, as shown in Figure 4.4. After treeing has progressed for some time, the stem and the larger branches grow hollow. Considerable discharges occur in these hollow spaces and form a special case for internal discharges; in contrast to the normal case, these discharges are very unstable as the trees may grow rapidly. No detectable discharges are seen for a long time during incubation; hours, weeks or even years can pass, depending on the applied voltage. Then treeing sets in, discharges become visible and grow so rapidly that they may cause breakdown in an extremely short period, sometimes in minutes or even seconds.

4.3. RECURRENCE OF DISCHARGES

The behaviour of internal discharges at AC voltage can be described using the well-known *a-b-c circuit* shown in Figure 4.5 [56]. The capacity of the cavity is represented by a capacitance C_c, which is shunted by a breakdown path. The capacity of the dielectric in series with the

cavity is represented by a capacitance C_b. The sound part of the dielectric is represented by capacitance C_a. The spark gap indicates that discharges may occur if the voltage is high enough. If a voltage V_a is applied to the insulation as shown in Figure 4.5, the voltage across the void is given initially by

$$V_c = \frac{C_c}{C_c + C_b} V_a \tag{4.1}$$

When this attains the value V^+, a discharge is ignited and the voltage collapses. The discharge extinguishes very rapidly (in about 10^{-7}s), the voltage across the void rises again, and discharges recur, as shown in Figure 4.6. The situation is complicated by the occurrence of a transient resistance and by chemical effects not taken into account by the simple circuit of Figure 4.5, but if the voltage is sufficiently high, discharges occur on every cycle. The walls of the void are eroded due to bombardment by charged particles so that the void lengthens, causing field intensification and, ultimately, breakdown.

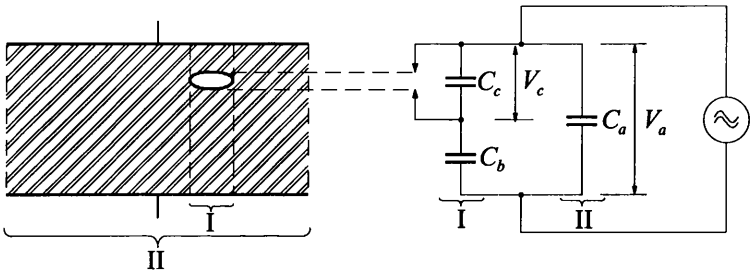


Figure 4.5: Void in a solid dielectric and equivalent circuit

Under an applied voltage V_a when V_c reaches the breakdown value V^+ of the gap, the cavity may break down. The sequence of breakdowns under a sinusoidal alternating voltage is shown in Figure 4.6. The dotted curve shows the voltage that would appear across the cavity if it did not break down. As V_c reaches the value V^+, a discharge takes place and the voltage drops down and the gap extinguishes. The voltage across the cavity then starts to increase again until it reaches V^+ when a new discharge occurs. Thus several discharges may take place during the rising part of the applied voltage. Similarly, on the decreasing part of the applied voltage, the cavity discharges as the voltage across it reaches V^-. In this way, groups of discharges originate from a single cavity and give rise to positive and negative current pulses on raise and decreasing part of the voltage, respectively.

The AC voltage across the sample at which discharges start to occur when the voltage is increased is called the inception voltage. If the voltage is decreased after discharges have been started, the voltage at which the discharges extinguish is usually lower than the inception voltage. Discharges may become intermittent if the cavity is asymmetrical. If a cavity exceeding a certain size, discharges may occur at different sites. The sequence of discharges is affected by transverse leakage of charge. The result is that small and large discharges succeed each other. A

sample usually has several cavities, and these in turn may contain different discharge sites; again, at one site several discharges occur recurrently. Consequently, if discharges are made visible on an oscilloscope, a complex picture will appear, consisting of constant, intermittent, and wandering discharges, separately or superimposed.

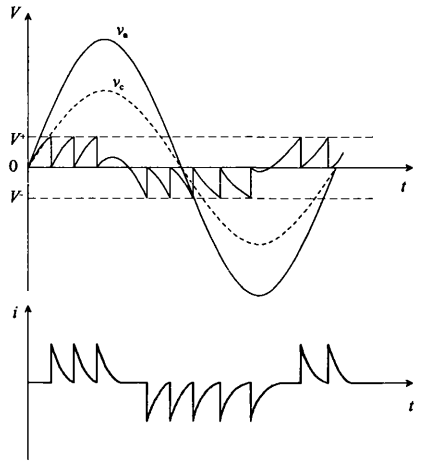


Figure 4.6: Sequence of cavity breakdown under AC voltage

In Figure 4.5, the faulty part of the dielectric corresponds to I, the sound part to II. The same representation can be given for surface discharges as shown in Figure 4.7. The surface that is covered by the discharge has a capacitance C_c to the electrode and a capacitance C_b through the insulation. The rest of dielectric is again represented by capacitance C_a.

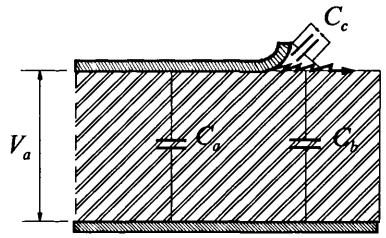


Figure 4.7: A dielectric circuit for surface discharges

Discharges by treeing can be represented in a similar way. The branches correspond to capacitance C_c, provided with a breakdown path. The breakdown paths have series capacitance C_b, which represent the sound part of the dielectric in series with a tree. The rest of the dielectric may be represented again by capacity C_a.

The behaviour of internal, surface and treeing discharges can be well described with *a-b-c* analogue circuit. Corona discharges can also be described with the *a-b-c circuit*, but the recurrence is different from other discharges. The analogue circuit is shown in Figure 4.8. The breakdown path at C_c is limited by space charges; these charges disappear quickly within milliseconds and determine the repetition rate of the corona discharges. If the AC voltage is slowly increased, corona discharges occur at first at the negative cycle of the sine wave only. In the beginning, one discharge peak per cycle is seen; a fast growing number of impulses is

shown if the voltage is raised fractionally. They are equal in size and their number increases about linearly with the applied voltage. At higher voltages, impulses also appear at the positive half cycle. These impulses are usually more irregular and are of a larger size.

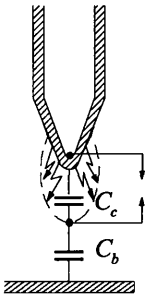


Figure 4.8: An analogue circuit for corona discharges

4.4. PD SOURCES IN TRANSFORMERS

PDs in transformers can occur at various positions in the windings, at leads, at sharp corners of edges of supporting parts, at defects or voids in the dielectric e.g. gas bubbles in oil or voids in cellulose material etc. Such PDs are even initiated by impurities as well as locally enhanced humidity values or a very high humidity content of the insulating material itself. Common sources of discharges in power transformers are [58]:

Delamination may occur when thinner pressboard sheets are glued together from thicker barriers.

Voids may occur differently. They may occur in glue. Furthermore, insufficient impregnation may give rise to voids. Sometimes (i.e., during commissioning) voids disappear when the oil absorbs gas and the cavity is filled with oil. Voids may also occur in bushings.

Bubbles may occur due to gas evolution from discharges, and evaporation of water droplets. Experience shows that a gas bubble in an open oil volume will be smashed into small bubbles that will quickly vanish when the first discharge occurs inside this bubble. Bubbles will therefore only exist at locations where they are supported mechanically by solid insulation (e.g., in wedges).

Discharge from metal parts having floating potential either to each other or to another part with a fixed potential. Free metallic particles may be left from the production process. Small particles will be carried by oil flow, while the largest will sink to the bottom. Free particles can easily adhere to surfaces (e.g., a winding).

Moisture may occur due to ageing and may also be introduced into the insulation during site erection. Moisture will contribute in several ways to discharge inception and extinction. During a heating cycle in service, moisture may be pressed out of the solid insulation. Due to

poor solubility in the oil, there will be super-saturation of the oil next to the cellulose surface resulting in water droplets and bubbles. Evaporation of the water will give rise to micro-bubbles, and hence, discharges. Increased moisture content in the cellulose will make it more conductive; therefore, cellulose fibres will start acting like metallic particles. Fibres sticking out from paper surfaces and fibres moving in the oil may initiate discharges. Increased moisture content will also increase the dielectric losses. During a cooling cycle, water may condense and be locally absorbed by pressboard. Pressboard has been observed to "puff up" due to the heat given off by dielectric losses and evaporation of water within the pressboard. As a consequence, internal discharges may become prevalent in the cavities within the pressboard.

5. PD MEASUREMENTS AS A DIAGNOSTIC TOOL FOR POWER TRANSFORMERS

5.1. INTRODUCTION

Partial discharges occur at locations where the electrical stress exceeds the limit of the insulating material. These limits depend on different parameters like the type of material, temperature, pressure, duration of stress, purity, etc. In any case, the inception of PDs at the given applied voltage demonstrates that at least locally, the withstand limit has been reached. Depending on the insulating material, the voltage level at which the PDs occur and the amount of discharges allow evaluation of the performance and quality of the insulation.

The detection and measurement of PDs present in the HV insulation is one of the most useful diagnostic tools for quality assurance testing during design, manufacturing and life assessment of any high voltage equipment. The technique for measuring and analysing PDs occurring in insulation systems normally includes the following steps: detection, measurement, recognition and evaluation, Figure 5.1.

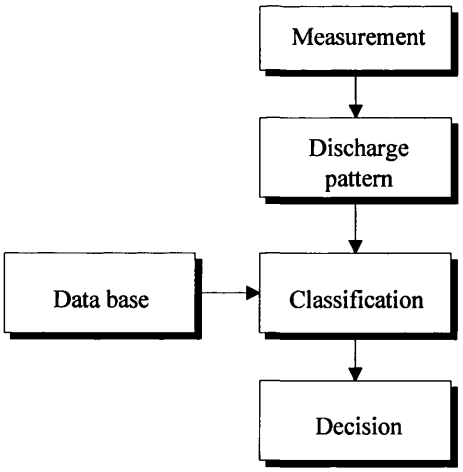


Figure 5.1: A general procedure of PD measurement and evaluation

5.2. DEFINITIONS OF QUANTITIES RELATED TO PARTIAL DISCHARGES [12]

Several quantities can be chosen to measure discharges:

q - *apparent charge* [pC] of a PD is that charge which, if injected instantaneously between the terminals of the test object, would momentarily change the voltage between its terminals by the same amount as the PD discharge itself

n - *pulse repetition rate* is the ratio between the total number of PD pulses recorded in a selected time interval and the duration of this time interval

N - *pulse repetition frequency* is a number of PD pulses per second, in the case of equidistant pulses

I - *average discharge current* [C/s or A] is the sum of the absolute values of the apparent charge magnitudes during a certain time interval, divided by this time interval

$$I = \frac{1}{t}\left[|q_1| + |q_2| + \ldots + |q_m|\right]$$

D - *quadratic rate* [C^2/s] is the sum of the squares of the apparent charges during a certain time interval, divided by this time interval

$$D = \frac{1}{t}\left[q_1^2 + q_1^2 + \ldots + q_m^2\right]$$

P - *discharge power* is the average power fed into the terminals of the test object due to PDs

$$P = \frac{1}{t}\left[q_1 u_1 + q_2 u_2 + \ldots + q_m u_m\right]$$

5.3. PD ACTIVITY DETECTION

The first step during the PD measurements is to ascertain as to whether the PD is present or not. PDs can react chemically with the surrounding atmosphere. They also create electromagnetic waves, acoustic waves and light. Therefore a PD detection system can make use of any of these phenomena.

5.3.1. Acoustic detection of PD activity

Acoustic detection of PDs within a transformer is based on the detection of the mechanical energy wave that propagates from the discharge site through the insulation. By placing a suitable sensor, the acoustic wave can be detected to generate information relevant to the PDs [9], [33], [60]. Acoustic techniques can be non-invasive, and therefore, they can be applied at most places on the transformer when it is in operation, and they are immune to electromagnetic noise.

In transformers, the use of piezo-electric acoustic emission sensors attached to the transformer tank wall has been the most favoured approach [8]-[10]. The PD site can be considered as a point source and hence the ultrasonic pulse associated with the discharge has a spherical wavefront. However, this wave will be attenuated and reflected differently by the acoustic interfaces/impedances, which it encounters in propagating from discharge site to the detection transducer(s). Typically, piezo-electric transducers for transformers will have a bandwidth between 20–350 kHz. In transformers, magnetostriction-induced core noise dominates the 50 kHz to 60 kHz region, therefore, sensors with resonance at either lower or higher frequency should be employed [107].

One of the advantages of the acoustic methods is that the site of a PD can be located by studying the phase delay or the amplitude attenuation of the acoustic waves picked up with more than three sensors located around the tank. The localisation techniques describes in the literature, see e.g. [9], [33], are generally based on triangulation. Although simple, this technique has some limitations. Complications arise due to the effects of signal attenuation especially in the presence of internal solid barriers such as core or other mechanical supports. More importantly, the location accuracy is poor in most cases due to the complex nature of the ultrasonic signals, which travel from the PD source to the sensor via various paths through the oil and along the transformer tank wall with different propagation velocities. From the composite signal picked up by the sensor, it is difficult to correctly identify the start of the wave-front associated with the different path signals so that the time delay can be measured and used for triangulation. Therefore, the authors in [8] and [59] use the three-transducer detector where the acoustic transducers are placed as close to each other as possible. This reduces the risk for generating varying paths and the localisation calculation is performed in two steps. Firstly, the average of the wave-front arrival times in all the sensors is used as if there were only one sensor. The direction of the source is then found using time differences between the sensors.

5.3.2. Optical detection of PD activity

Problems associated with externally mounted piezoelectric sensors, such as corruption of signal from environmental noise and difficulties to locate the exact site of PDs due to multi-path of the acoustic wave, makes it desirable to have sensors that can operate inside a transformer reliably and pick–up clean PD–induced acoustic signals. Optical fibre-based sensors have been shown to be attractive to measure a wide range of physical and chemical parameters because of many advantages, including small size, light weight, high sensitivity and frequency response, and immunity to electromagnetic interference. These attributes make optical fibre sensors an excellent candidate for transformer PD detection.

Blackburn et al. [61] have investigated the use of a fibre optic detection system within the power transformer tank. With this arrangement, the acoustic signal based PD propagates by a direct route and is detected before being corrupted by the composite oil/steel signal. Detection is based on a principle that the ultrasonic pressure wave impinging on the fibre causes mechanical stress in the fibre core. This changes both the effective fibre length and its refractive index. Consequently, the phase of the light traversing the fibre will also be affected. It should be noted that this approach has not yet realised its potential but work continues to improve sensitivity and reduce interference.

The fibre optic sensor uses a silica diaphragm and a single mode optical fibre encapsulated in a silica glass tube to form a sensing element is presented in [62]. It is used for inside detection of PDs in power transformers. Compared with the conventional acoustic sensors, the fibre optic sensor has the advantages of non-electrically conducting, high frequency response, immunity to the electromagnetic interference, small size, and the capability of multiplexing more than one sensor in a single fibre.

5.3.3. Chemical detection of PD activity

PD activity invariably results in changes to the chemical composition of the particular phases involved. In power transformers, PD activity can be inferred by the chemical detection of modifications in the transformer oil. These manifest themselves in fault-generated gases in the oil and paper insulation degradation bi-products. The current chemical approach detects PDs in transformers by taking oil samples from the transformer (see Chapter 3.3.2.). Problems associated with chemical methods are due to the fact that there can be a long time delay between the initiation of a PD source and the evolution of enough gas to be detectable. The further limitation is that it is generally not possible to determine the exact location of the detected PD source.

5.3.4. Electrical measurement of PD activity

A variety of electrical techniques are in current use to assess the condition of power transformer insulation [12], [64]-[66]. Techniques include the use of narrowband and broadband PD detectors, which provide data on each individual discharge event, integrated measurements of loss within the insulating system associated with PD activity, and various aerial techniques. PD signals can be decoupled with Rogowski coil [64]-[65], capacitive dividers [65]-[66], tap bushings [65], or antennas [13]. Sometimes it is difficult to separate internal and external discharges; therefore, directional coupling techniques are suitable [64]-[65]. They determine the power flow direction of the PD signals, thereby enabling a distinction between signals from inside and outside the transformer. It is well known that individual discharge magnitudes can

indicate the relative size of individual degradation sites and that the magnitude/phase relationship of these pulses on the power cycle and the way these vary with applied voltage and time can provide valuable data on the nature and form of prevalent degradation [63]. The application of digital technology to this type of measurement technique has enhanced its potential further by providing significantly improved means of acquiring, storing and processing the PD data in an objective, repeatable and fast manner.

Traditionally, electric PD measurements are performed off-line using a capacitive decoupling and a detection unit. IEC publication 60270 provides a general standard for conventional PD measurements [12]. The apparent charge is used as the quantity to be measured and the test circuit and measuring system should be calibrated.

The purpose of conventional PD measurement is the detection and relative quantification of electrical discharges (and thus possible defects) in the insulation system of the test object. The most common PD measurement circuit is shown in Figure 5.2, where C_t is the capacitance of the test object (be C_a, C_b and C_c in Figure 4.5), i_{TE} the PD induced current pulse, C_c the connecting cable capacitance, C_k the coupling capacitor and Z_m the measurement impedance.

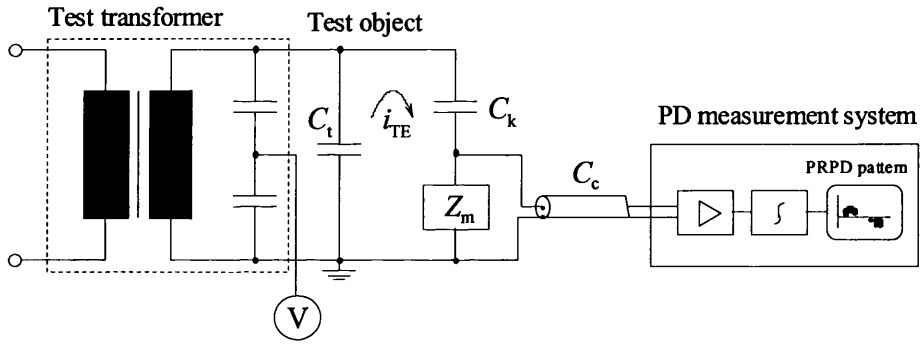


Figure 5.2: Conventional PD measurement circuit

A PD in the test object causes a fast and therefore high frequency pulse i_{TE} in the test circuit. Since the test transformer presents relatively high impedance to high frequency signals, the bulk of the current pulse flows through the coupling capacitor and the measurement impedance. The resulting voltage signal is transmitted via a coaxial cable to a PD measurement device. The typical frequency range of this type of measurement is 40 kHz to 1 MHz. Simpler PD measurement apparatus presents the signal on a peak reading meter, often combined with an oscillogram to indicate the position of current pulse relative to the voltage frequency phase. Newer types of measurement systems record both the amplitude and the phase of each PD pulse relative to the test voltage; thus they enable phase resolved PD analysis (see Chapter 5.4.1.1).

With regard to the frequency range, two types of systems exist: wideband and narrowband PD measuring system. The recommended values for a wideband system according to [12] are:

$$30 \text{ kHz} < f_1 \leq 100 \text{ kHz};$$

$$f_2 \leq 500 \text{ kHz}$$

$$100 \text{ kHz} \leq \Delta f \leq 500 \text{ kHz};$$

where f_1 and f_2 are the lower and upper frequencies, and Δf is the bandwidth. The response of these instruments to a PD current pulse is in general a well-damped oscillation. The apparent charge q as well as the polarity of the PD current can be determined from this response, and the pulse resolution time T_r is small and is typically 5 to 20 μs.

Narrowband measurements are characterised by a midband (centre) frequency f_m and its bandwidth Δf. Recommended values for Δf and f_m are

$$9 \text{ kHz} \leq \Delta f \leq 30 \text{ kHz}$$

$$50 \text{ kHz} \leq f_m \leq 1 \text{ MHz}$$

They enable noise suppression by virtue of the selection of a suitable centre frequency. However, it is necessary to choose a centre frequency of a few MHz for adequate noise suppression. This is not in accordance with the IEC 60270 where the centre frequency is limited to 1 MHz, except if the frequency spectrum is almost constant up to the chosen centre frequency. Otherwise, a measurement in the high frequency range exclusively allows only a statement about the existence of PDs, but not about their apparent charge. The response of a narrowband system to a PD current pulse is a transient oscillation with the negative and positive peak values of its envelope proportional to the apparent charge q, independent of the polarity of this charge. The pulse resolution time T_r is typically about 80 μs.

5.4. PD PATTERN ANALYSIS TECHNIQUES

Each defect has its own particular degradation mechanism; therefore, it is important to know the correlation between discharge patterns and the type of defect. In this respect, the time discharge quantities can be divided into [67]:

1. Basic quantities, which are quantities observed during one voltage cycle:
   - Discharge magnitude q
   - Ignition voltage V_i
   - Position of the discharge relating to the phase angle φ_i of the test voltage
   - Numbers of discharges for each half period of the voltage cycle N_q.
2. Deduced quantities, which are integrated values of basic quantities from the first group, observed throughout several voltage cycles (> 50 cycles). These quantities can be analysed as a function of time and as a function of the phase angle.

Classification is based on recognition. There are two basic possibilities for recognizing discharges: phase related and time resolved recognition [68]. In *phase related recognition* the patterns that occur in the 50 Hz sinewave are studied. Each impulse on the display is specified

by its magnitude q and its phase angle φ. The resulting pattern of all impulses is a characteristic of the type of discharge. In *time resolved recognition,* the true shape of the charge displacement in the defect, e.g. cavity, is shown at a nanosecond scale. From these shapes, comprehensive conclusions can be drawn: the physical state of the discharge site, the type of the discharge, and even the ageing process which takes place in or at the dielectric.

5.4.1. Phase related PD recognition

A PD activity is a sequence of PD pulses $q_s(t_s)$ with their apparent charge q_s and their time of occurrence t_s, as shown in Figure 5.3. Since the PD pulse occurrence is related to the phase of the applied voltage, PD acquisition systems also record the magnitude of the test voltage v. Thus a PD activity of N pulses can be acquired with all the information considered relevant, as a sequence of pulses $q_s(t_s,v(t_s))$, for each s (s=1,2,...,N) over a measuring time t_m. Such a PD record with the essential information about the development of the PD activity is called a *phase resolved pulse sequence* (PRPS) measurement, which can be analysed by different procedures.

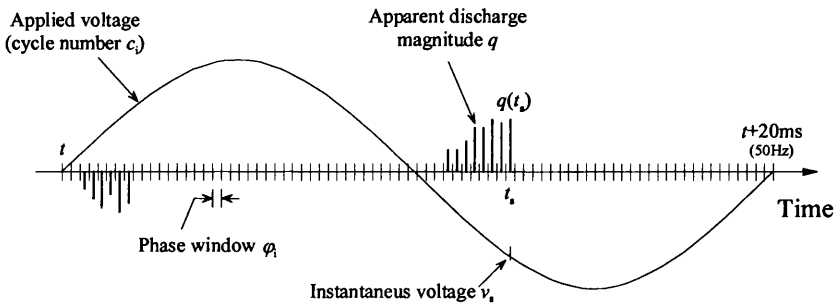


Figure 5.3: A PD pulse sequence and relevant parameters

One of the main targets in PD identification is to relate the statistical characteristics of PD distribution to properties of defect at which the discharge occurs and eventually to identify the type of defect. Phase related recognition consists in general of three stages [68] as shown in Figure 5.4. In the first stage, *statistical distributions* are derived from discharge pattern. The most commonly used distributions are [68]-[73]:

- $H_N(\varphi)$ the pulse count distribution
- $H_{qn}(\varphi)$ the mean pulse height distribution
- $H_{qmax}(\varphi)$ the maximum pulse height distribution
- $H_{\Sigma q}(\varphi)$ the average discharge current distribution.

These distributions have characteristic shapes, which vary with the type of the defect. Since the discharge activity usually occurs on both halves of the AC voltage, these distributions should be calculated separately for positive and negative half of the sinwave.

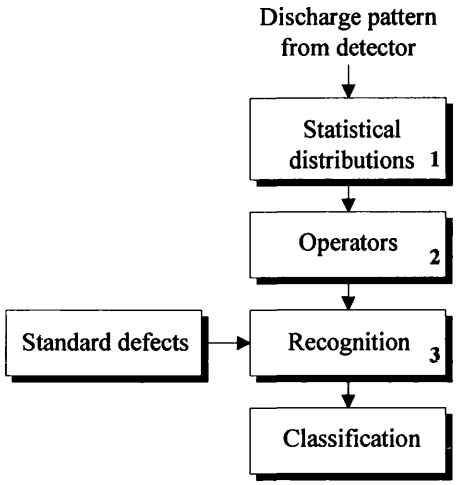


Figure 5.4: Three stages of phase related recognition

The second stage *operators* (i.e. mathematical descriptions or parameters) are applied to distributions from stage one to describe their characteristic shape. The most commonly used operators are [69]:

- The Skewness (S_k) represents the asymmetry of the distribution.
- The Kurtosis (K_u) represents the sharpness of the distribution.
- The cross correlation factor (*cc*) indicates the difference in the shape of the distributions in the positive and negative half cycle.
- The number of peaks (P_e) enables discrimination between distributions with single top and distributions with several tops (mixed distribution).
- The Discharge factor (Q) describes the difference in the mean discharge level in the negative and positive distribution.
- The Modified cross correlation factor (*mcc*) equals the product of Q and *cc*.

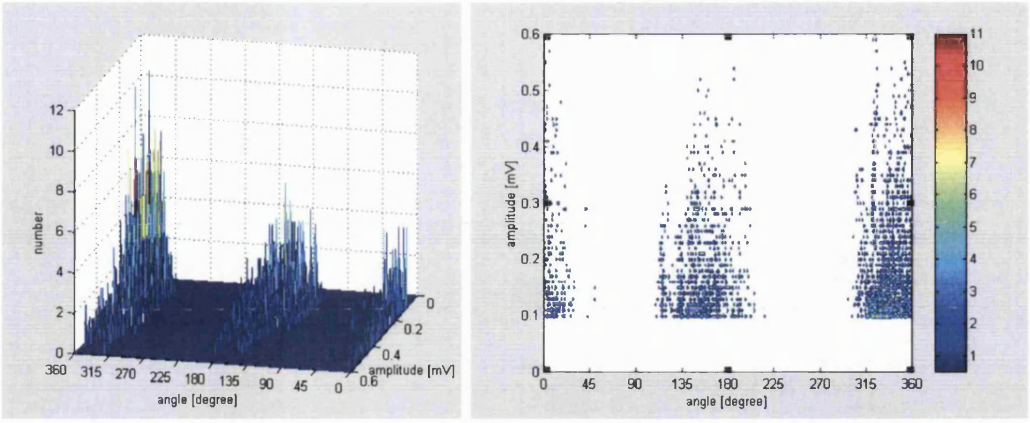
Operators applied to the distributions serve as a fingerprint of a certain defect. A fingerprint could be compared to the other prints in order to characterise a discharge. The third stage is *recognition* where operators are compared to those of known discharges.

Major PD patterns and visualization tools are presented below.

5.4.1.1. Phase resolved partial discharge (PRPD) pattern

The so-called φ-q-N pattern $H(q, \varphi)$, which is also called a PRPD pattern, is a widely spread tool for PD expert evaluation [69]-[73]. An example of this graph (Figure 5.5, left), in the shape of a 3D histogram, features the number of PD pulses with a certain position φ and the apparent charge q of a PD event. This plot can be changed into a contour plot, the so-called 'topogram', where in the third dimension; the number of pulses is presented by a colour or grey

value (Figure 5.5, right). A φ-q-N pattern can be considered as an integrated plot where the sequential characteristic of the PD pulses with respect to time is lost.



3D pattern Topogram

Figure 5.5: φ-q-N pattern

5.4.1.2. Δu pattern

The Δu pattern in Figure 5.6, evaluates the frequency of the voltage differences of a sequence of PD pulses showing the voltage range necessary to evoke consecutive PD pulses [80]. The Δu-values are calculated from the test voltage values at the time of occurrence of consecutive pulses (see also Figure 5.7). However, it must be emphasised that the PD re-ignition activity depends strongly on the voltage gradient and not on the absolute voltage itself.

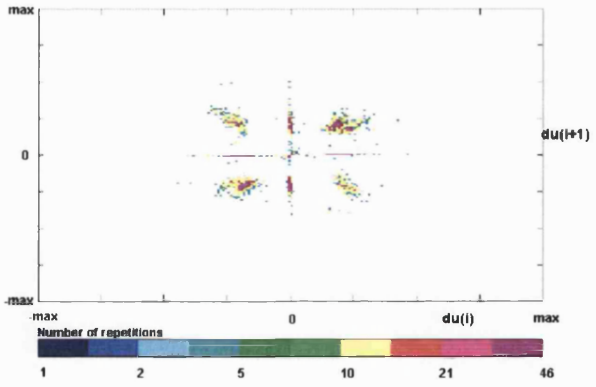


Figure 5.6: Example of Δu pattern H(Δui+1/Δui)

5.4.1.3. Δu/Δφ Pattern (pulse sequence pattern)

In the Δu pattern only the voltage difference is considered while in the $\Delta u/\Delta\varphi$ pattern, the voltage gradient $m=\Delta u$ of consecutive pulses, (as shown in Figure 5.8), is evaluated in a histogram $H(m_{i+1},m_i)$. Equation (5.1) describes the slope m_i between two consecutive PD pulses

with number i-1 and i, where α_i is the inclination of the secant that contains the points (φ_{i-1}, $u(\varphi_{i-1})$) and (φ_i, $u(\varphi_i)$). φ_i is the phase position and $u(\varphi_i)$ the instantaneous voltage value of PD pulse i.

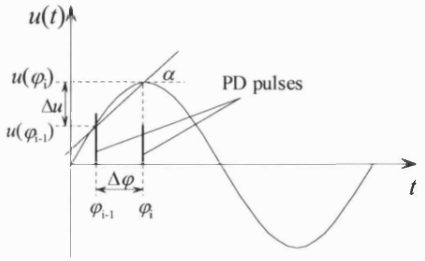


Figure 5.7: Voltage gradient and voltage difference of two consecutive pulses

$$m_i = \left.\frac{\Delta u}{\Delta \varphi}\right|_i = \frac{u(\varphi_i) - u(\varphi_{i-1})}{\varphi_i - \varphi_{i-1}} = \tan \alpha_i \quad \forall\, i = 2,\ldots,M \tag{5.1}$$

The slope m_i approximates the voltage gradient which is necessary to excite a consecutive PD pulse i.

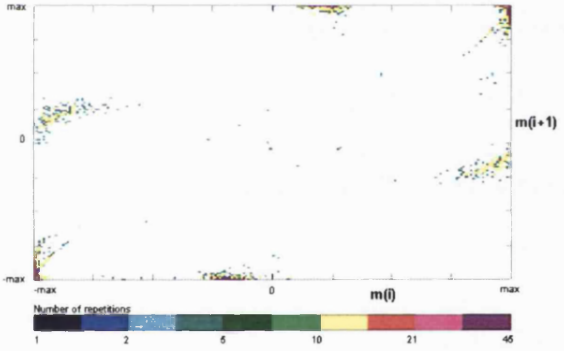


Figure 5.8: Voltage gradient and voltage difference of two consecutive pulses

5.4.2. Time resolved recognition

In order to preserve the shape of the discharge pulse and relate it to the physics of the discharge process, a time resolved detection system is needed. A measuring system with a bandwidth higher than 50 MHz is able to classify impulses due to their different signal shapes caused by different PD faults or PD shaped interferences [78], [81]. The signal shape recognition determines the following PD pulse parameters of each detected PD pulse:

- apparent charge Q,
- 50% pulse duration T,
- PD pulse amplitude $\hat{U}$,

- pulse polarity,
- phase angle of the test voltage φ,
- amplitude of the test voltage $\hat{U}_p$.

Figure 5.9 illustrates a typical PD pulse and the associated measured PD parameters. The apparent charge Q is simply determined by integration between zero crossing t_1 and t_2 of the PD signal:

$$Q = \int_{t_1}^{t_2} i(t)dt \tag{5.2}$$

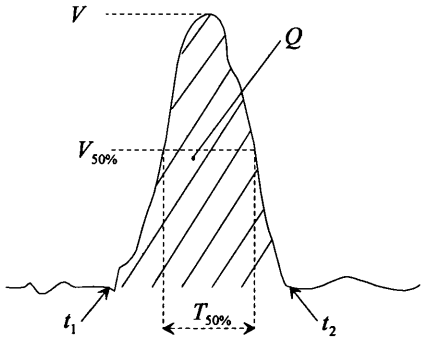


Figure 5.9: Pulse parameters in the time domain

The main advantage of wide-band PD measurement is that the PD signal contains more information to describe the pulse shape.

5.5. CLASSIFICATION AND DECISION

The aim of classification is to assign a label to a PD pattern of unknown origin from previously collected patterns with known labels, such as corona, floating particle, etc. This means that a data base of previously collected patterns must be available. In the past, a number of methods have been used for classification purposes:

- conventional classifiers [74],
- hidden Markov models [75],
- fuzzy classifiers [76],
- artificial neural networks [77].

On the basis of the classification result, i.e., when a PD pattern of unknown origin has been identified such as internal discharge, a decision must be made; for example, the transformer should be disconnected from the network and detailed inspection of its condition should be carried out. Of course, the decision is based on knowledge of the potential danger posed by different defects derived from past experience.

6. POWER TRANSFORMER NON-INVASIVE PD MEASUREMENTS – FIRST TRIAL

6.1. INTRODUCTION

The use of remote radiometric measurements for identifying PD behaviour in high-voltage plant has not been investigated in any great detail in the past. This technique is potentially advantageous in field measurements because it requires no galvanic connections to the transformer, thus allowing measurements to be made from operational transformers without undue complications. The thesis describes the first attempt to make measurements of this type. Radio frequency interference (RFI) measurements of PDs were performed on a 1000 MVA, 400/275/33 kV power transformer, produced by Hackbridge & Hewittic at Neilston 400 kV substation in Scotland.

It was envisaged that the measurements performed at the Neilston substation would help towards answering the following questions:

1. At what distance can a PD be detected?
2. Can the PD position be detected?
3. What is the sensitivity of a PD measurement?
4. Can different insulation faults be recognised by studying the shape of the individual PD pulses?
5. Is it possible to detect the existence of two or more PD defects at the same time?
6. How can external interferences be detected and eliminated?

6.2. TEST SETUP

The 400/275/33 kV, 1000MVA power transformer at the Neilston substation is not connected to the power network. Since the transformer has had known discharges, it was energized to activate them.

A mobile diesel generator unit and a step-up transformer 0.433/33 kV were connected to the compensating (delta) winding of the unit under test as shown in Figures 6.1 and 6.2, respectively.

Figure 6.1: A diesel generator unit and a step up transformer 0.433/33 kV (left) and a 400/275/33 kV transformer (right) under test for PD-diagnosis

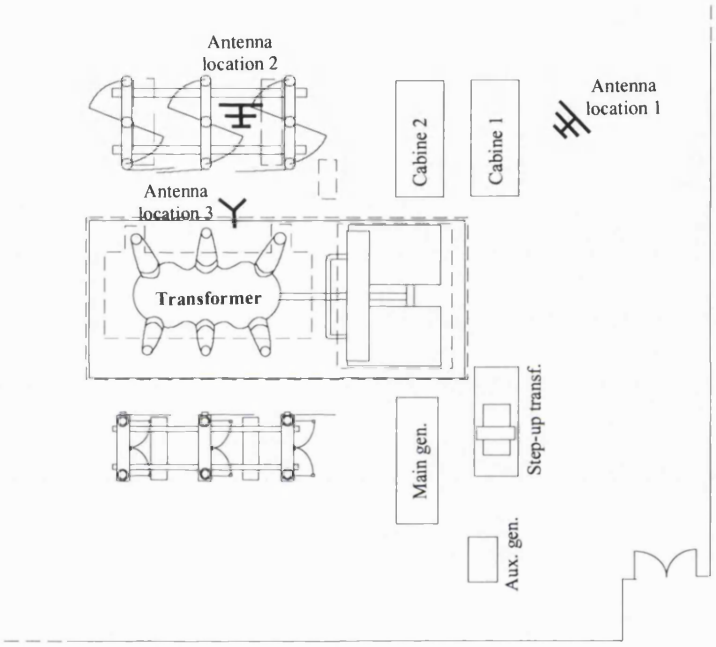


Figure 6.2: Ground plan of tested transformer unit with generator unit and step-up transformer and antenna locations.

6.3. PARTIAL DISCHARGE ACQUISITION SYSTEM

The PD-measuring system (as shown in Figure 6.3) consists of the following components: different types of antennas (Table 6.1), receiver, digital phosphor oscilloscope, and laptop.

A directional Yagi antenna was used for measurements at both locations 1 and 2; a monopole antenna was placed at location 1, while a helical antenna was located next to the transformer at location 3, as shown in Figure 6.2. Antennas at location 1 were connected with the receiver by semi-rigid cables while aerials at locations 2 and 3 were linked with the receiver via 50 Ω coax cable.

Besides PDs, the oscilloscope also captured applied AC voltage. The digital phosphor oscilloscope has 2 channels, maximum sample rate 2.5 GS/s and bandwidth of 300 MHz. The oscilloscope was connected with the laptop via a GPIB communication module. The scope was set to record 20 ms (i.e. one cycle) of data, rather than being triggered on individual PD pulses. The GPIB interface allowed the acquisition system to transfer data to the computer for storage, approximately every 2 seconds.

Table 6.1: Antennas and their frequency range used for PD measurements

Antenna	Frequency range	Radiation pattern
Helical	300 … 500 MHz	omnidirectional
Yagi	70 … 130 MHz	directional
Monopole	30 … 70 MHz	omnidirectional

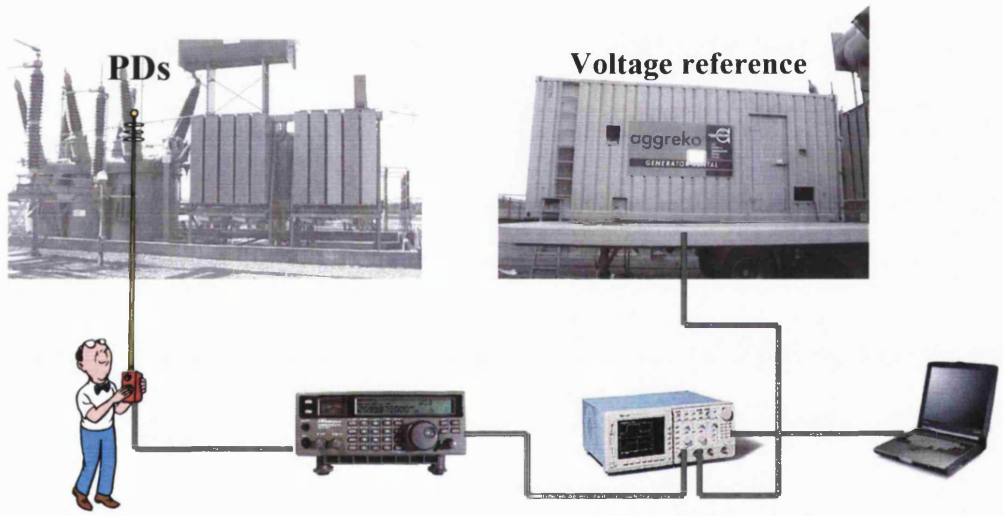


Figure 6.3: Schematic representation of the PD measuring system: set-up of principal components

6.4. RESULTS

First of all, the centre frequency of reception had to be determined. Frequencies were scanned between 2 and 116 MHz and the best results were obtained at a frequency of 61.5 MHz. Therefore, the reception frequency of all measurement results presented in this section are centred at 61.5 MHz. Note that this centre frequency was slightly outside the range of the Yagi (frequency range of 70 - 130 MHz), and far away from the helical tuned frequency range of 300-500 MHz, however, measurements using these two antennas were still possible. With

regard to the reference [82], the received signals were down sized to 10.5 MHz intermediate frequency. The reason to perform down converting a signal is a desire to perform more complex analysis of the signal. At lower frequencies, it is possible to digitise the signal to a higher degree of precision.

Figures 6.4 and 6.5 show PD and background noise signals measured by monopole antenna and Yagi antenna, respectively. It must be emphasized that the background noise was measured before the transformer was energised. Therefore, blue signals in Figures 6.4 and 6.5 contain beside PDs also the background noise. To better illustrate the pattern generated by the PDs and background noise, Figures 6.4 and 6.5 use a span of 100 ms, as opposed to a span of 20 ms that is used in the subsequent figures.

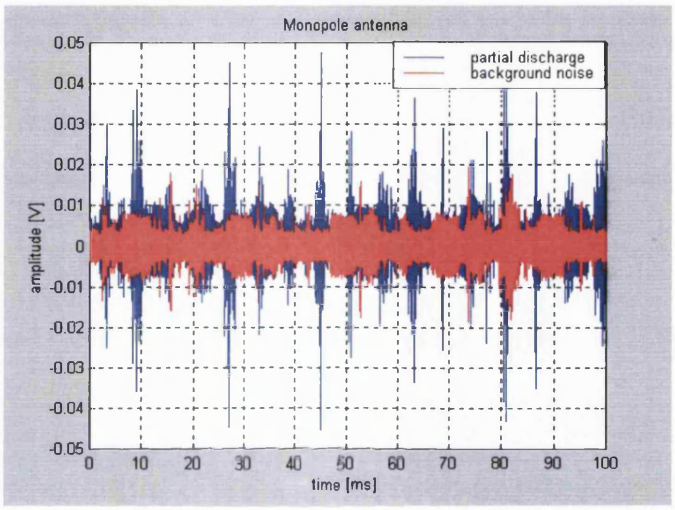


Figure 6.4: Captured PD signals and background noise by monopole antenna

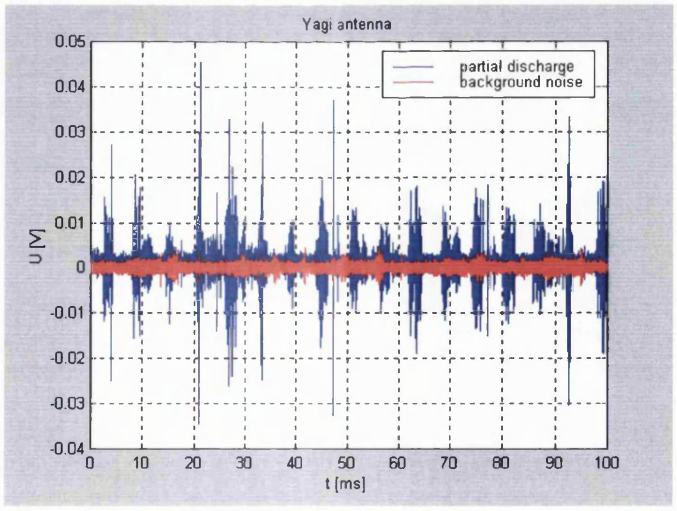


Figure 6.5: Captured PD signals and background noise by Yagi antenna placed beside the cabin

Figure 6.6 is an actual screen capture of the oscilloscope with a time window of 20 ms, i.e. one period of an AC cycle. The applied voltage was 100% of the rated voltage. The monopole antenna was used to capture the PD signals. Unfortunately, the bandwidth of the system was insufficient to examine separate pulses (as shown in detail in Figure 6.6) and thus an investigation of a single pulse was not possible.

For ease in comparison of the results, the signals were aligned so that for each cycle, the voltage signal starts at the zero value. Since the background noise was in all cases lower than 0.015 V, all discharges below this value were ignored.

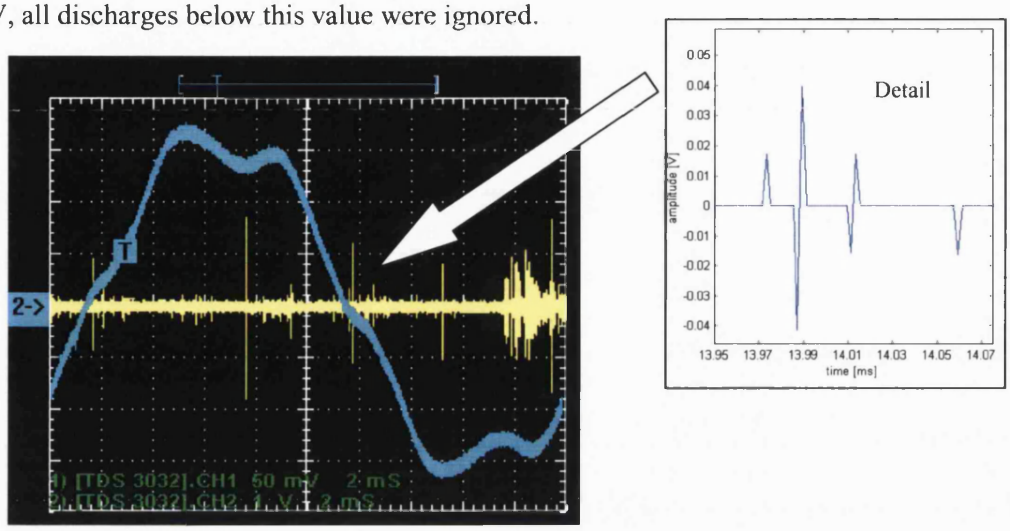


Figure 6.6: Measured PDs signal by monopole antenna together with the voltage signal

The transformer under test was energised, without load, through its tertiary windings by an auxiliary 33kV transformer. This in turn was energised by a generator. The test set up allowed variation of the excitation voltage up to almost 120% of the rated working voltage. The frequency of the test voltage was variable between 50 and 60Hz. Since the transformer was not loaded, the only current was magnetizing current. Consequently, the line voltage was distorted as seen in the following figures.

Figures 6.7 through 6.13 present one cycle of signals measured with different antennas, placed at different locations. Position of the discharge relating to the phase angle φ_i of the line voltage (u) is given. The applied voltage was 100% of the rated voltage and the frequency was 50 Hz. Since the Yagi antenna is directional, it was hoped that this could be directed to the particular phase of the transformer, thus identifying the phase that produced the discharge. To this end, discharges were captured with the Yagi antenna in both vertical and horizontal polarizations. Unfortunately, the results did not show any significant variation. The likely cause of this is due to the Yagi operating frequency range of 70 – 130 MHz being higher than the 61.5 MHz optimum centre frequency identified during the tests; moreover, at lower frequencies, the Yagi loses its directional capability.

The abbreviations in figures are V-voltage, V_r-rated voltage and f- frequency.

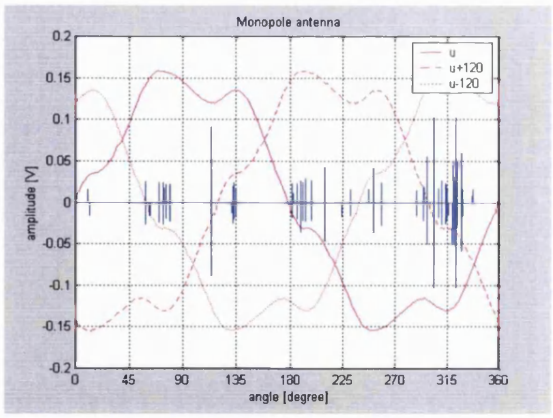


Figure 6.7: PD measurement by monopole antenna, $V = 100\%V_r, f = 50$ Hz, location 1

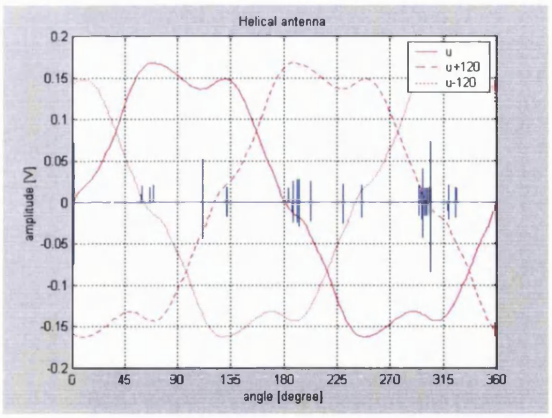


Figure 6.8: PD measurement by helical antenna, $V = 100\%V_r, f = 50$ Hz, location 3

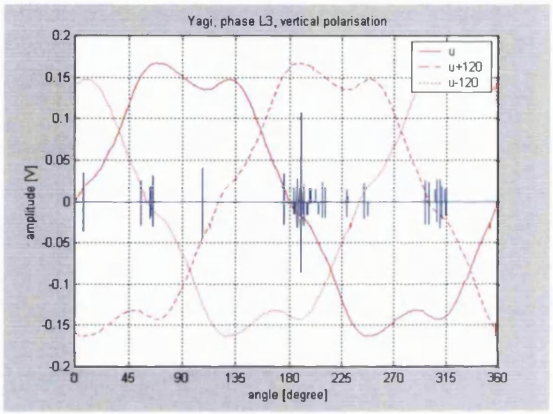


Figure 6.9: PD measurement by Yagi antenna in phase L3, vertical polarisation, $V = 100\%V_r, f = 50$ Hz, location 2

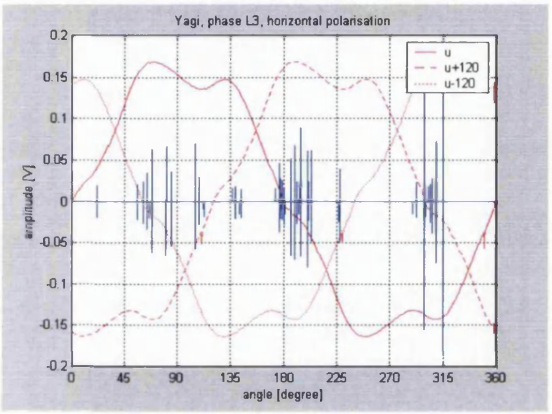


Figure 6.10: PD measurement by Yagi antenna in phase L3, horizontal polarisation, $V = 100\%V_r, f = 50$ Hz, location 2

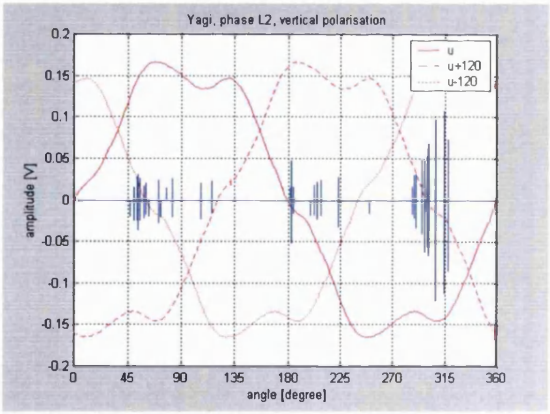


Figure 6.11: PD measurement by Yagi antenna in phase L2, vertical polarisation, $V = 100\%V_r, f = 50$ Hz, location 2

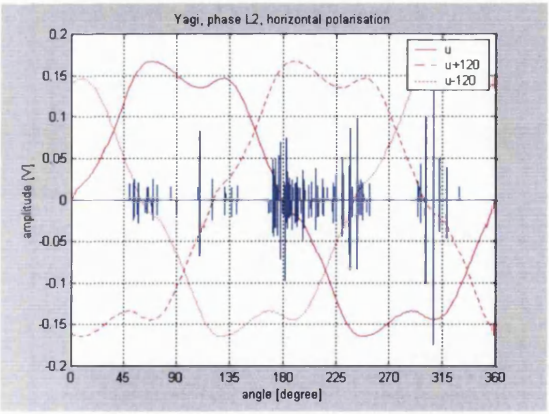


Figure 6.12: PD measurement by Yagi antenna in phase L2, horizontal polarisation, $V = 100\%Vr, f = 50$ Hz, location 2

Figures 6.14 through 6.17 also illustrate one cycle for signals measured with different antennas. The position of the discharge related to the phase angle φ_i of the applied voltage is given again, but this time the frequency of the voltage was 60 Hz in order to minimize PDs of the other equipment in the substation. In Figure 6.14, the applied voltage was 85% of the rated voltage and as expected it is seen that the occurrence of discharges is not so high in amplitude or intensity as compared to Figure 6.15 where the applied voltage was 115% of the rated voltage. In both cases, measurements were performed with the Yagi antenna placed at location 1 and polarised vertically. Figures 6.16 and 6.17 give results obtained with monopole and helical antennas which are omnidirectional. The applied voltage was 115% of the rated voltage.

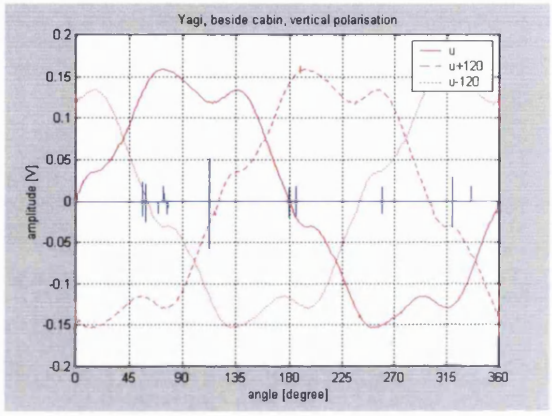


Figure 6.13: PD measurement by Yagi antenna beside cabin, vertical polarisation, $V = 100\%V_r, f = 50$ Hz, location 1

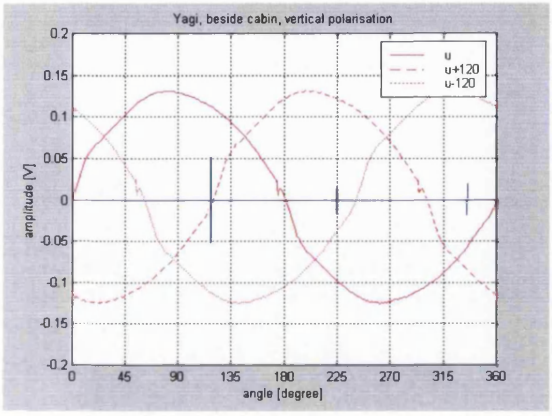


Figure 6.14: PD measurement by Yagi antenna beside cabin, vertical polarisation, $V = 85\%V_r, f = 60$ Hz, location 1

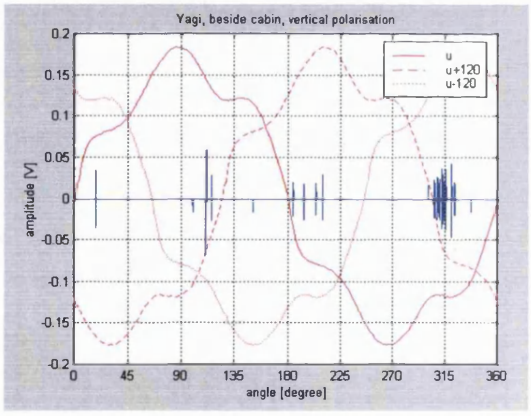


Figure 6.15: PD measurement by Yagi antenna beside cabin, vertical polarisation, $V = 115\%V_r, f = 60$ Hz, location 1

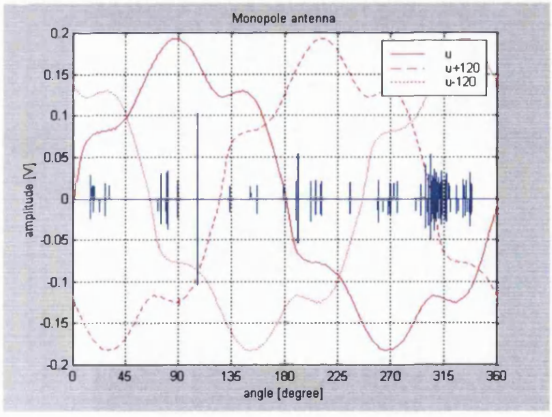


Figure 6.16: PD measurement by monopole antenna, $V = 115\%V_r, f = 60$ Hz, location 1

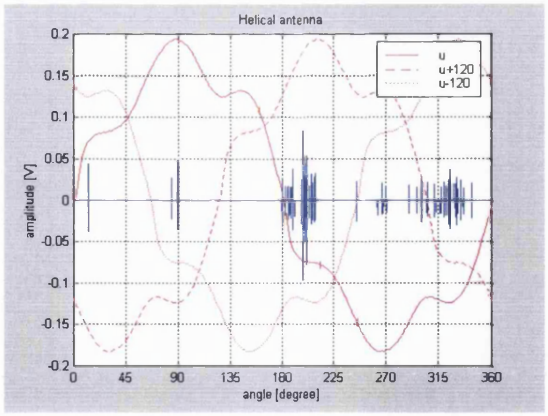


Figure 6.17: PD measurement by helical antenna,
$V = 115\%V_r, f = 60$ Hz, location 3

In Figures 6.18 through to 6.22, 3D and 2D PD fingerprints are illustrated (see Chapter 5.4.1.1). They give the relationship between the PD magnitude and its intensity as a function of phase angle. Additionally, these figures show the results for 50 cycles of recorded data. On the left side of each figure is a 3D pattern presentation and on the right side is a topogram. Figure 6.18 presents PD fingerprints obtained by the Yagi antenna at location 2. The applied voltage was 100% of the rated voltage with a frequency of 50 Hz. All other figures give results with a frequency of 60 Hz. Figures 6.19 and 6.20 illustrate measurements obtained with the Yagi antenna placed at location 1 with the voltage 85% and 115% of rated voltages, respectively.

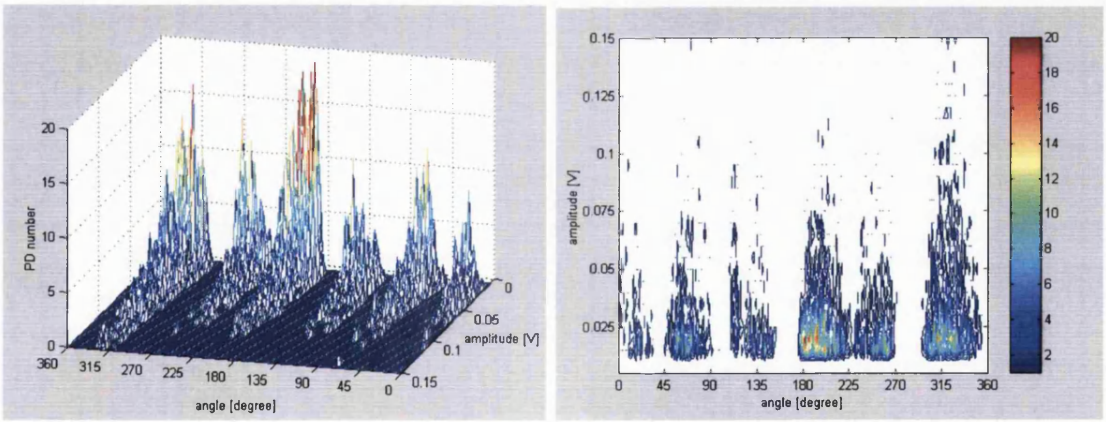


Figure 6.18: PD measurement by Yagi antenna, location 2, $V = 100\%V_r, f = 50$ Hz
Observation throughout 50 voltage cycles

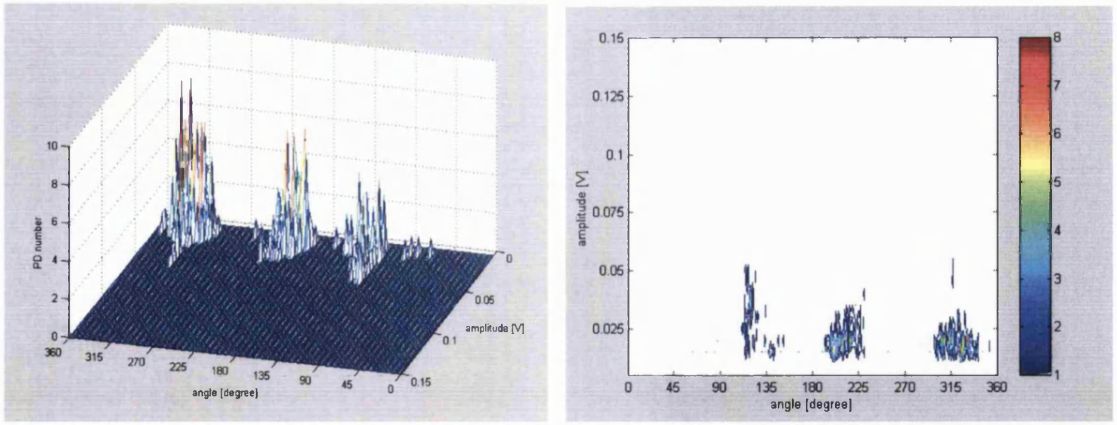


Figure 6.19: PD measurement by Yagi antenna, location 1, $V = 85\%V_r, f = 60$ Hz
Observation throughout 50 voltage cycles

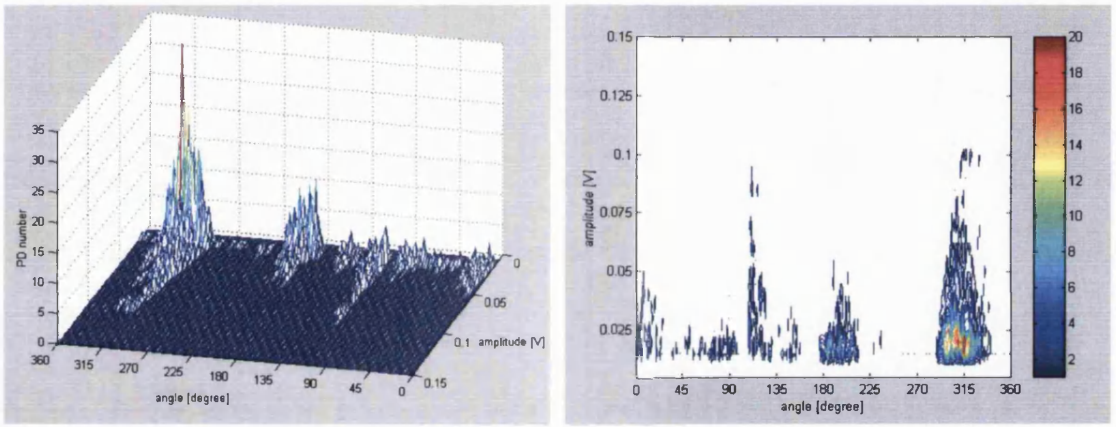


Figure 6.20: PD measurement by Yagi antenna, location 1, $V = 115\%V_r, f = 60$ Hz
Observation throughout 50 voltage cycles

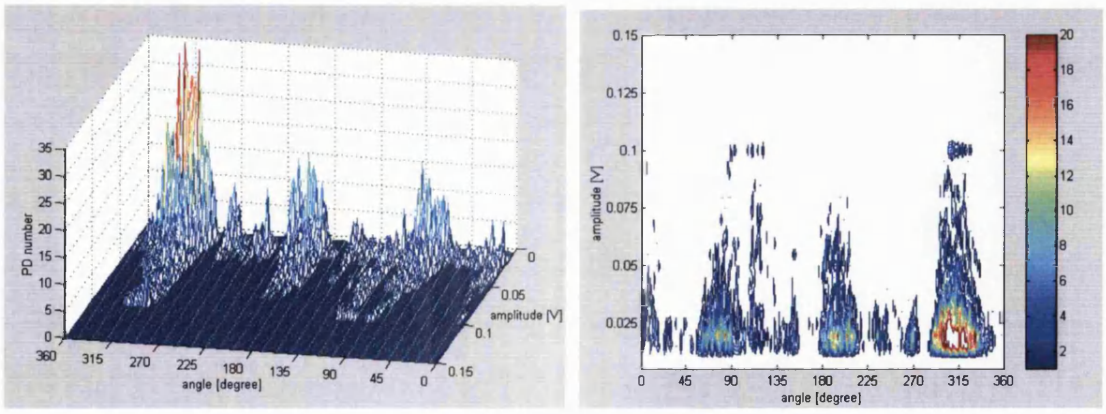


Figure 6.21: PD measurement by monopole antenna, location 1,V = 115%Vr, f = 60 Hz
Observation throughout 50 voltage cycles

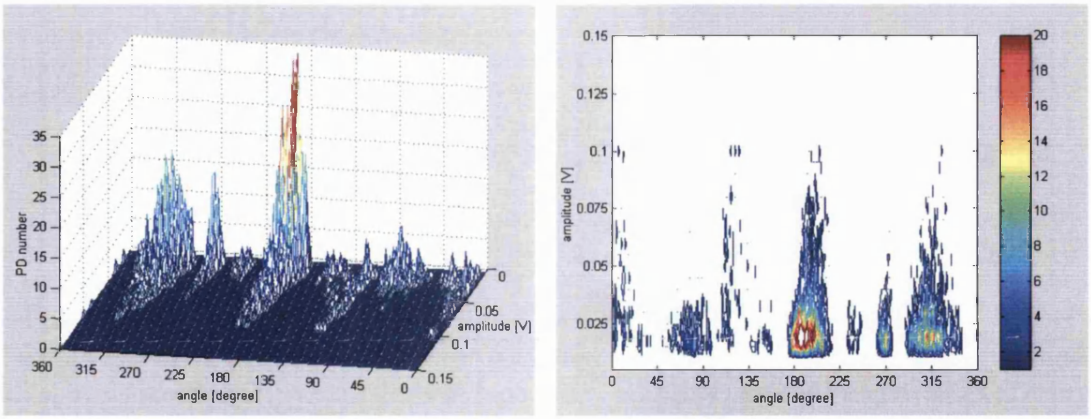


Figure 6.22: PD measurement by helical antenna, location 3, V = 115%Vr, f = 60 Hz
Observation throughout 50 voltage cycles

6.5. CONCLUSION AND DIRECTIONS

The results show a general correlation between discharge pulse intensity and density as a function of the applied voltage. It is conjectured that the majority of the pulses recorded emanated from surface discharges on or around the transformer HV bushings. Due to the distorted reference voltage – which displayed multiple peaks – it was not obvious to which phase each group of pulses could be attributed; this problem was further compounded by the lack of antenna directionality. The results also suggest that one phase produced more corona noise than the others, as shown in Figure 6.21. It can be also seen that discharges were captured in all phases (approximately 120° difference between the main PD concentrations, as shown in Figure 6.21).

From the results obtained, it has not been possible to distinguish between benign corona effects and unwanted internal PDs. The topogram in Figure 6.21 shows the possibility of multi-source PD patterns. However, it is likely that classification of the two effects will require greater resolution of waveforms than was possible during these tests. Additionally, vital information would be provided if the recording system had the ability to discriminate between discharges produced on individual phases.

For future tests, the following recommendations are made:

- Wider bandwidth recording equipment should be used.
- Multiple, simultaneous antenna recordings should be made in order that the location of the discharge can be determined.
- Each recorded PD signal should be expanded in order to analyse its shape.
- Recordings should be made over many cycles of data.

7. WIDE BAND MEASUREMENTS

7.1. INTRODUCTION

This chapter describes the PD investigation carried out on the 400/275 kV autotransformer SGT3A at the Northfleet substation. The investigation based on the RFI measurements for detection and identification of discharges occurring in power transformers is presented. The measurement system consists of wideband antennas, an oscilloscope and a laptop. The well-established KEMA probe [13], [83] for monitoring PDs was mounted in the oil valve at the bottom of the transformer tank. The signals from the probe serve as a reference signal for the measurements obtained by RFI measurements in order to separate signals caused by discharges occurring outside and inside the transformer. The measured data is presented as phase resolved and time resolved discharge patterns.

7.2. SET UP OF THE DETECTION SYSTEM

With regard to the references [83] and [84], the current pulse of PDs is very short in time, the order of nanoseconds. Consequently electromagnetic waves with frequencies up to 1 GHz are emitted. Therefore antennas should be capable detecting ultra high frequencies. The PD-measuring system consists of TV Yagi (frequency range 470-850 MHz) and helical (frequency range 300-500 MHz) antennas, digital phosphor oscilloscope, and laptop. Figure 7.1 shows a ground plan of the transformer and arrangement of Yagi (left) and helical (right) antennas. Since the Yagi aerials are directional, they were pointed at the bushings in phases A (Yagi 2) and C (Yagi 1) on the low voltage side (LV) and to phase C (Yagi 3) on the high voltage (HV) side. The letter K denotes the position of the KEMA UHV antenna mounted on the oil valve at the bottom of the transformer tank.

Antennas on the low voltage side were connected with the oscilloscope by 33m long semi rigid cables while antennas on the high voltage side by 33m long satellite type cables.

A Tektronix TDS 7104 Digital Phosphor Scope can sample four channels synchronously at a sampling rate of 2.5 GS/s and has an analogue bandwidth of 1 GHz. The channel 1 captured

data from one of the antennas, channel 2 from the KEMA probe and channels 3 and 4 served for capturing magnitude (V) and the slope (dV/dt) of the reference voltage.

In order to preserve the shape of the discharge pulse, the wide band measurements were performed. The oscilloscope enables acquiring data with frames. The so-called "FastFrame" is an acquisition mode that lets one capture many records in a larger record and then view and measure each record individually. The setting of FastFrame was 100 frames of 5000 samples each. The measurement window was 2μs with $\Delta t = 0.4\ 10^{-9}$ s. Figure 7.2 shows an example of one acquisition of pulses measured on the 275 kV side close to phase C by the helical antenna. The left figure shows all 100 measurements and the figure on the right shows a detail of one of the 100 frames.

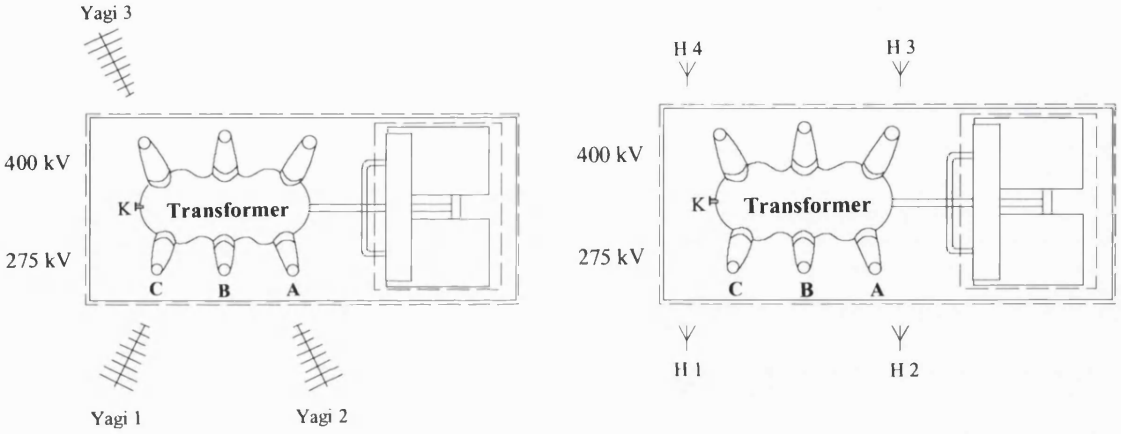


Figure 7.1: Ground plan of tested transformer and arrangement of antennas

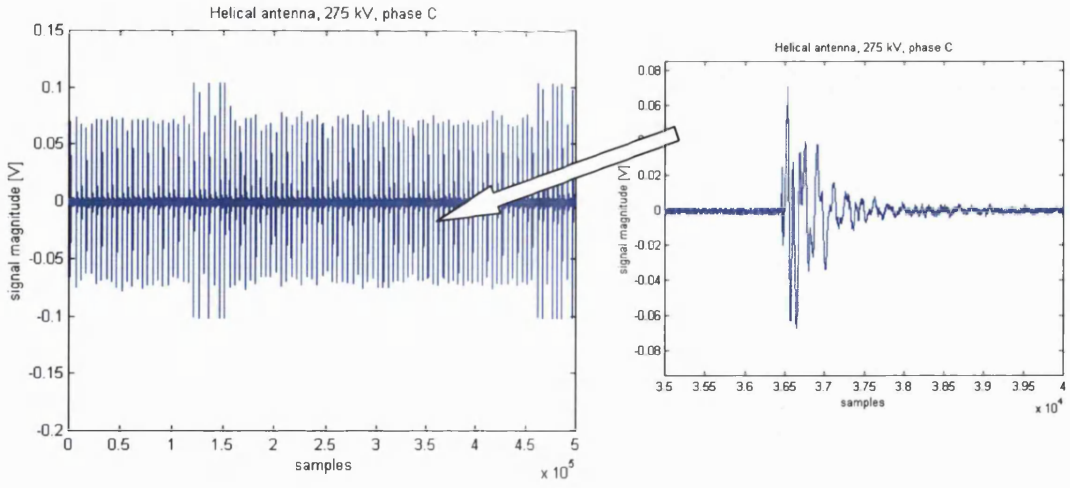


Figure 7.2: One measurement of 100 pulses captured on the transformer

7.3. TYPICAL MEASURED SIGNALS

Captured signals are presented in the time and frequency domains and some of the 2D statistical results are also included. As mentioned before, signals obtained from KEMA probe serve as a reference for signals obtained by the aerials.

Figures 7.3 to 7.5 show background noise signals and their frequency spectra measured by Yagi, helical antenna and the KEMA probe, respectively. The signals were not recorder at the same time. From the measured signals, it can be seen that the noise is in general caused by radio station (88-110 MHz) and TV signals (430-650 MHz). The magnitude of the signals is smaller than 8 mV in the case of Yagi and helical antenna and below 2.5 mV in the case of the KEMA probe.

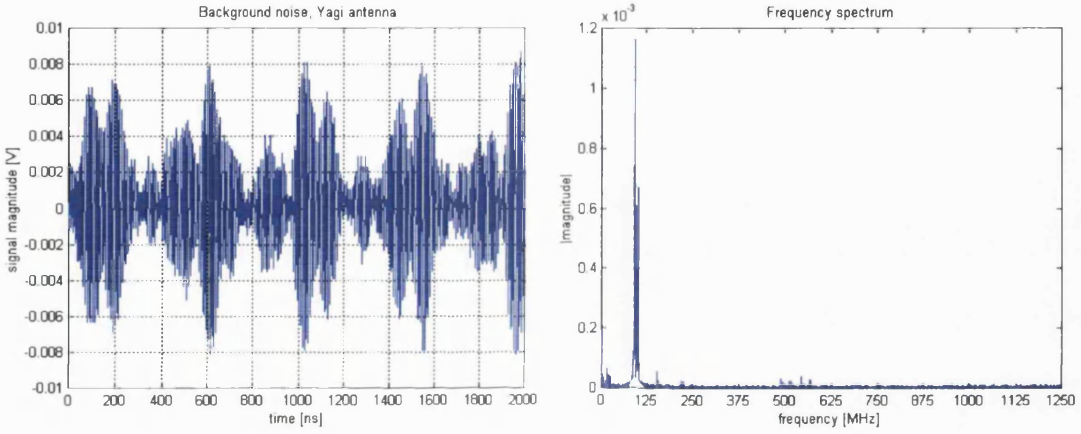


Figure 7.3: The background noise of the signal and its frequency content measured by Yagi aerial

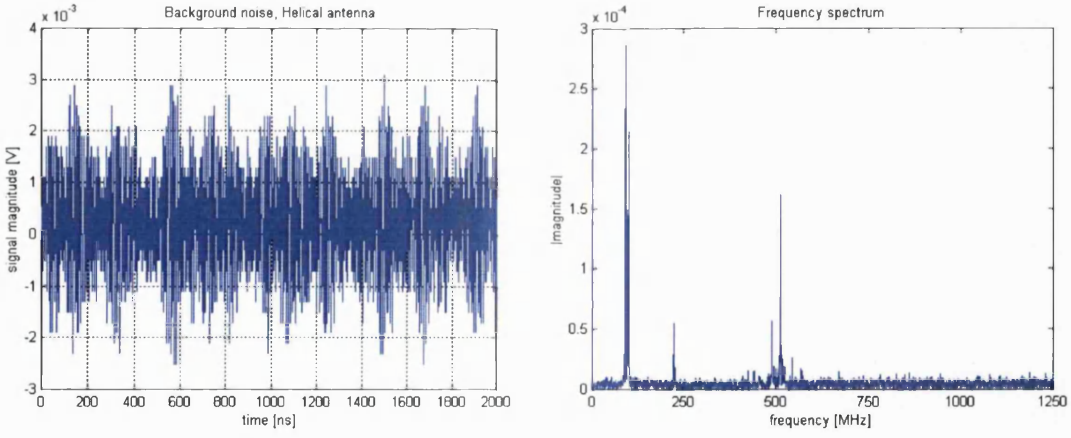


Figure 7.4: The background noise of the signal and its frequency content measured by helical antenna

In general, signals can be divided into three to six representative groups. Each group usually contains two signals (mirror image), one occurring on positive half of the reference

voltage and one occurring on the negative half of the reference voltage; both have nearly the same frequency spectrum.

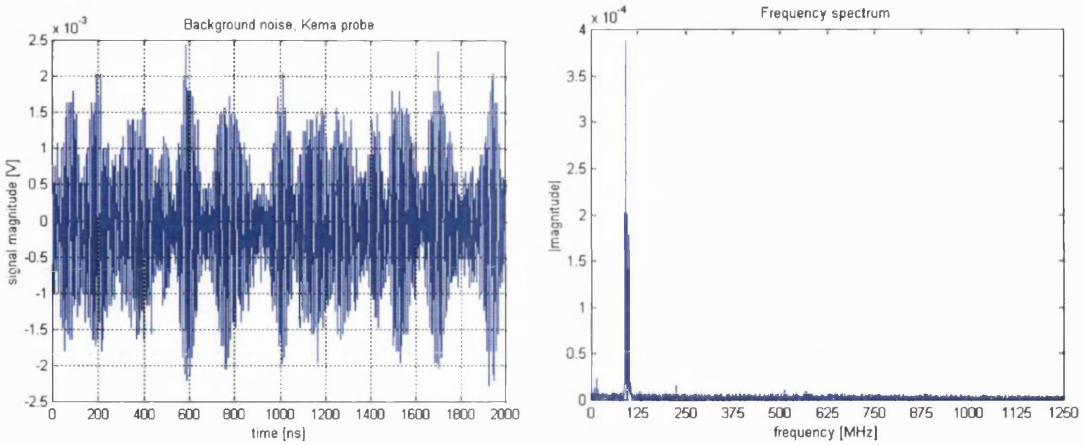


Figure 7.5: The background noise of the signal and its frequency content measured by the KEMA probe

Herein, only typical measured signals for Yagi1 and helical antenna H1 are given. For all other locations, typical signals can be found in Appendix.

7.3.1. Typical signals obtained by Yagi 1

Figures 7.6 to 7.8 show three main groups of signals and their frequency spectra measured by Yagi 1, pointing at phase C on the 275 kV side of the transformer (left) and signals detected by the KEMA probe (right). The signals were triggered by Yagi signals. The arrival time of the impulse at the Yagi and the KEMA probe is slightly different due to different lengths of connecting cables.

The shapes of the 1st and the 2nd group of signals are quite similar; the 2nd group of signals detected by Yagi are smaller in magnitude. The 3rd group of signals, Figure 7.8, were only detected by Yagi antenna and the signals are small in magnitude while the KEMA probe could not detect a "serious" signal, except radio frequency signals. The third group of signals belongs to discharges occurring on the antenna.

Blue signals are those occurring on the positive half of the reference voltage while red signals present signals occurring on the negative half.

Figures 7.9 to 7.11 show the number of pulses with regard to the voltage phase angle. Figure 7.9 illustrates the distribution of all groups of discharges, together. Figure 7.10 presents distribution of the number of all discharges with the green colour and with the black colour discharges of the 1st and 2nd group while Figure 7.11 shows the situation for the 3rd group of discharges separated for the positive (left side) and negative (right side) voltage cycle.

Discharges of the first two groups occur between 350° and 40° and between 165° and 215° while discharges of the 3rd group taking place between 40°-50° and 130°-140° on the positive cycle and between 220°-230° and 310°-320° on the negative cycle.

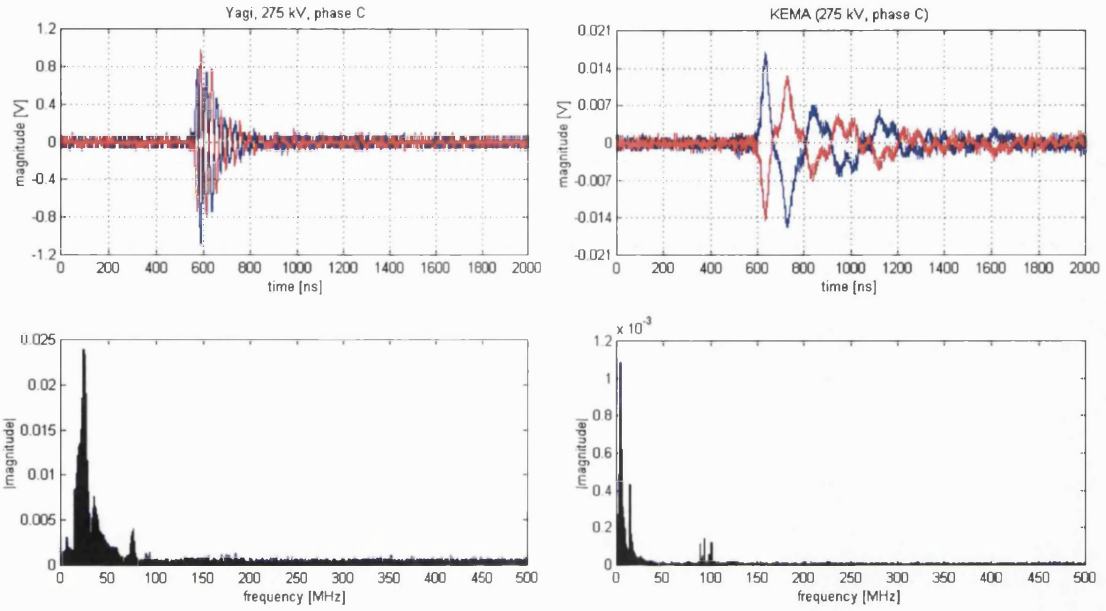


Figure 7.6: 1st group of signals measured by Yagi 1 and the KEMA probe; triggered on Yagi signals

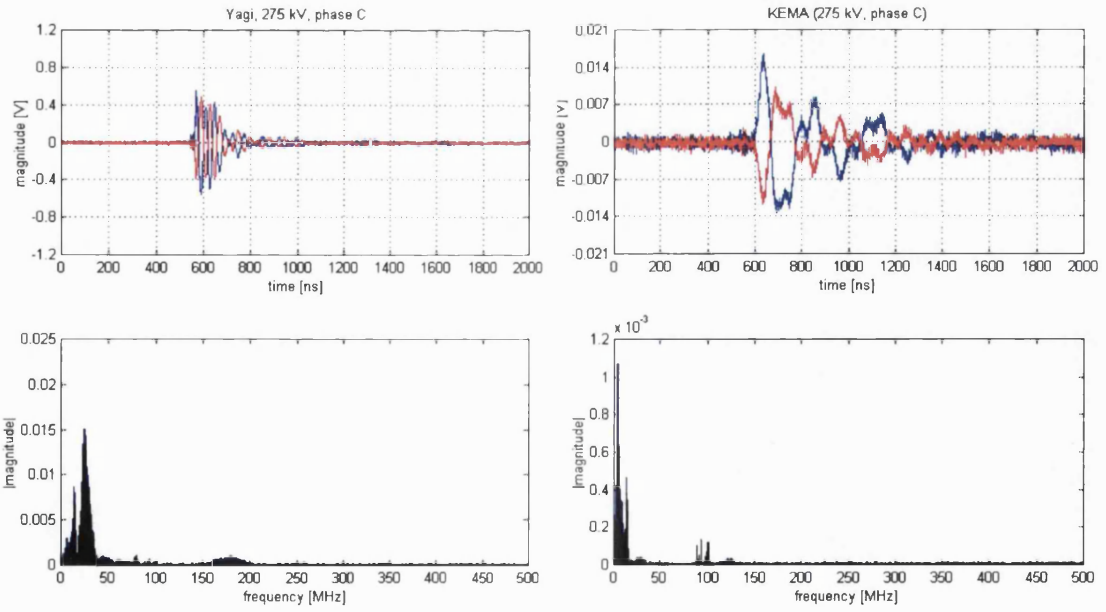


Figure 7.7: 2nd group of signals measured by Yagi 1 and the KEMA probe; triggered on Yagi signals

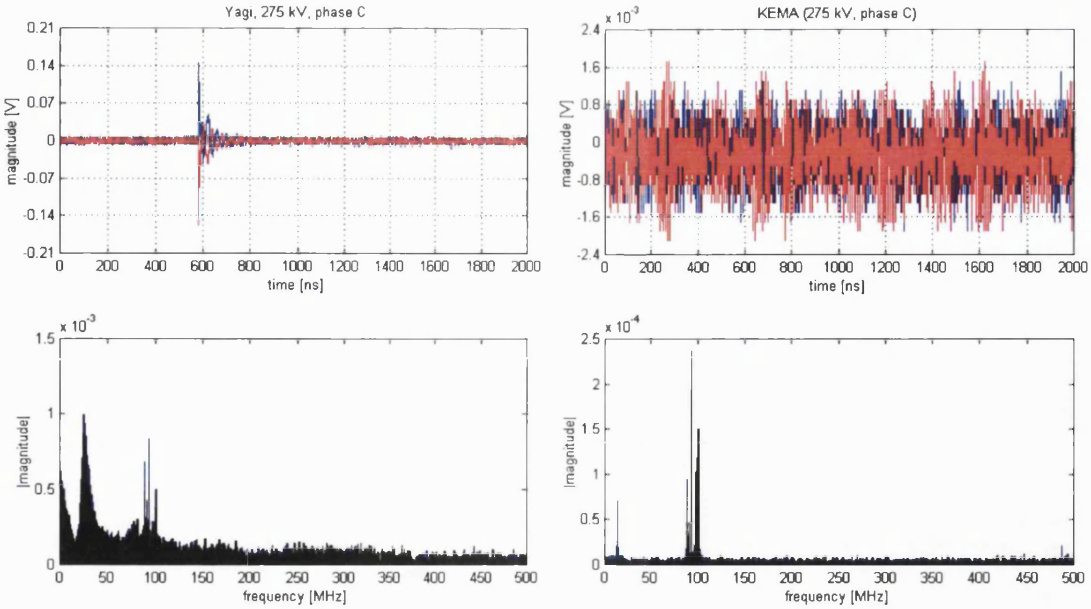


Figure 7.8: 3rd group of signals measured by Yagi 1 and the KEMA probe; triggered on Yagi signals

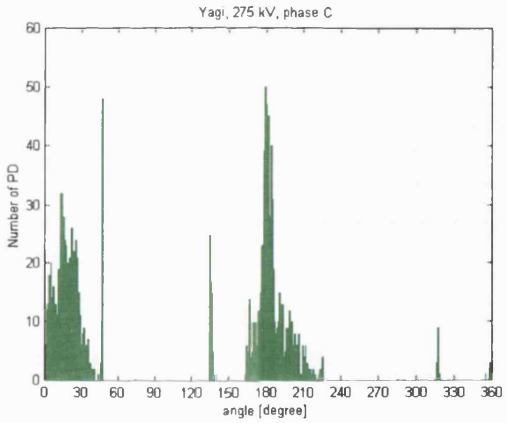


Figure 7.9: Distribution of the number of pulses regarding the voltage phase angle for signals measured by Yagi 1

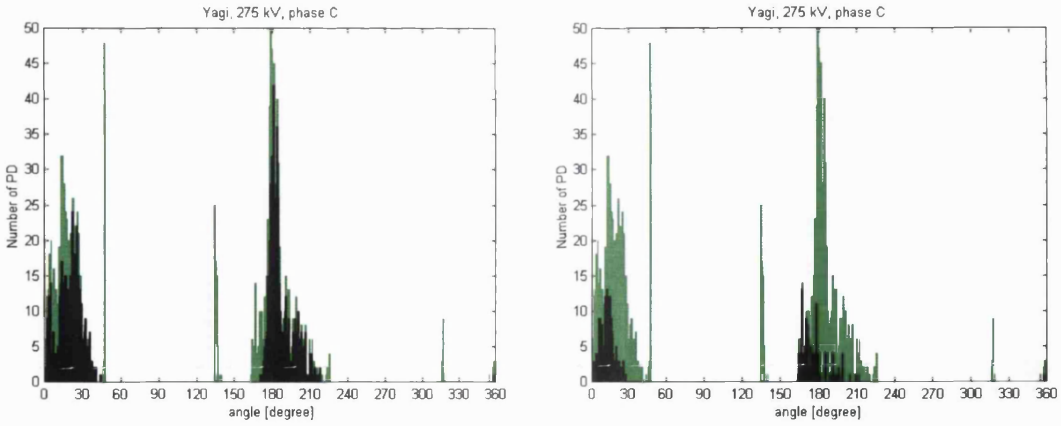


Figure 7.10: Distribution of the number of pulses regarding the voltage phase angle for the 1st (left graph) and the 2nd (right graph) group of signals

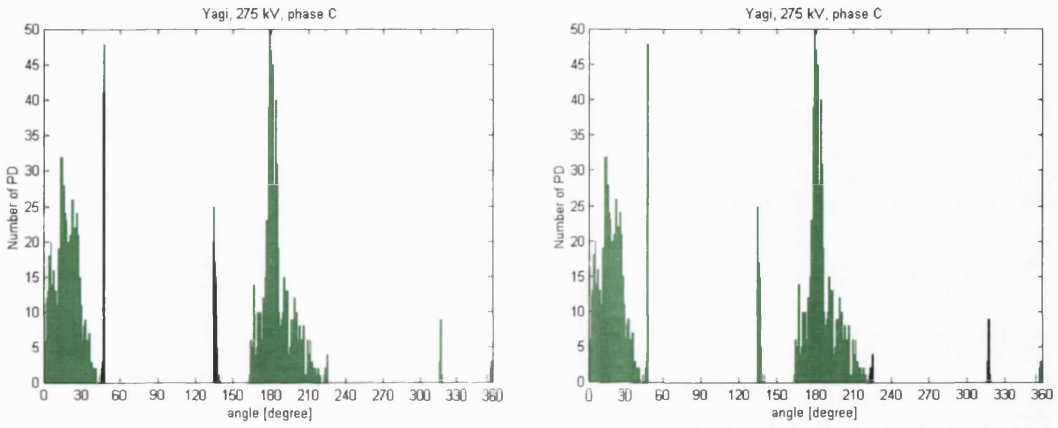


Figure 7.11: Distribution of the number of pulses regarding the voltage phase angle for the 3rd group of signals; left side – positive cycle and right side –negative cycle

7.3.2. Typical signals obtained by H1

Figures 7.12 to 7.15 show four main groups of signals and their frequency spectra measured by the helical antenna H1, placed close to phase C on the 275 kV side of the transformer (left) and signals detected by the KEMA probe (right). The signals were triggered by helical antenna signals.

The shapes of the first two groups of signals are quite similar. Signals contain dominant frequencies below 100 MHz, typical for discharges in the air. The 3rd group of signals is small in magnitude and show the characteristic of very fast pulses and also the frequency spectrum contains frequencies higher than 200 MHz, which could possibly indicate the presence of discharges in oil [83], [84].

Figures 7.16 to 7.18 show the number of pulses with regard to the voltage phase angle. The distribution of the first two groups is between 0° - 50° and 175° - 220°. Figure 7.18 (left side)

shows the situation for the 3rd group of discharges occurring between 25° - 50°, 130° - 140°, 190° - 220° and at 315°.

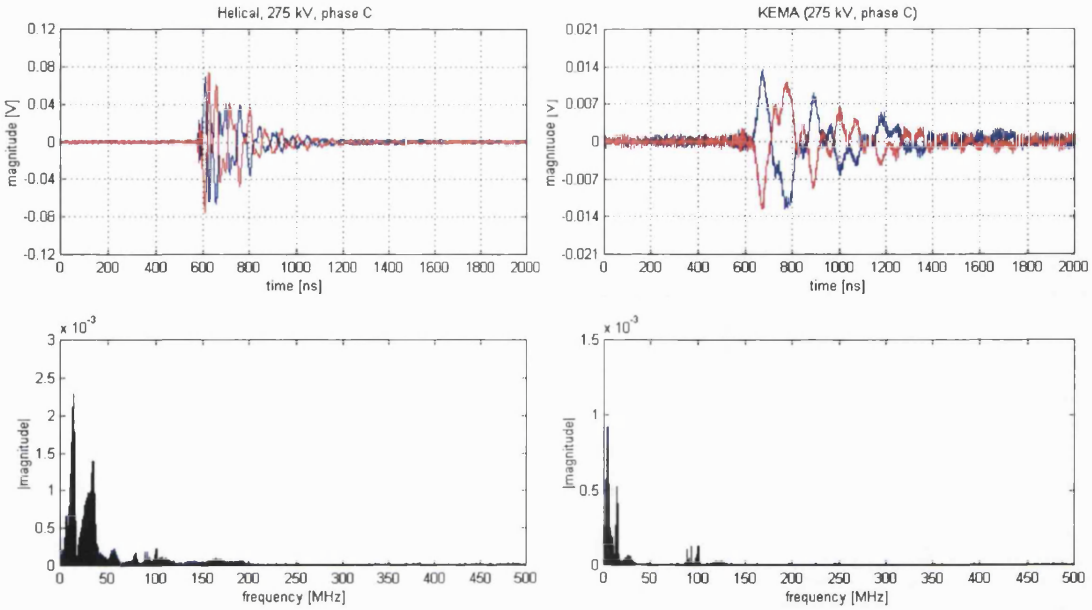


Figure 7.12: 1st group of signals measured by the helical antenna H1 and the KEMA probe; triggered on helical antenna signals

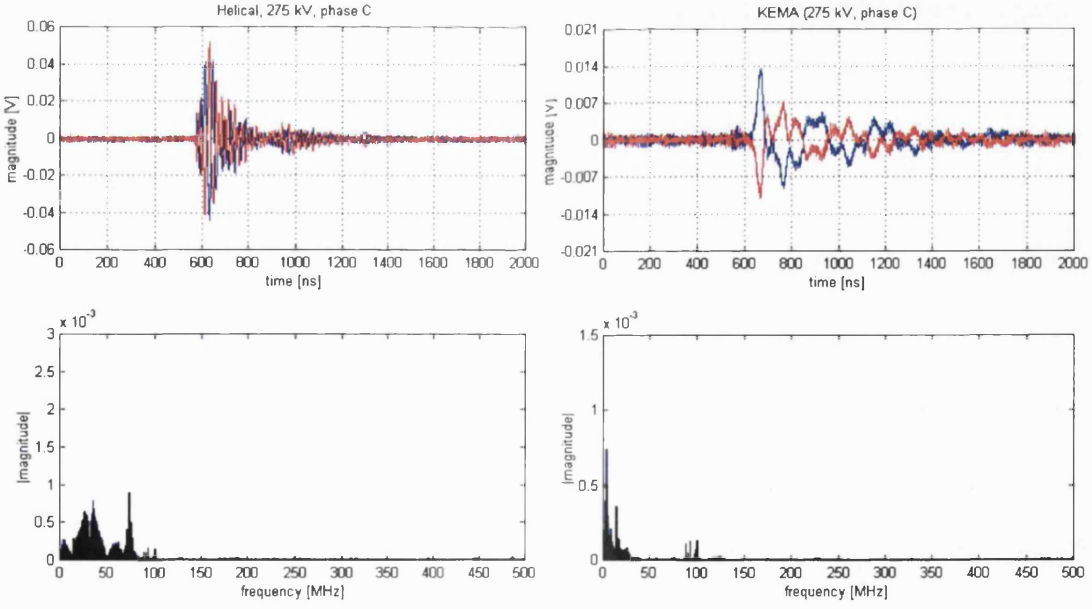


Figure 7.13: 2nd group of signals measured by the helical antenna H1 and the KEMA probe; triggered on helical antenna signals

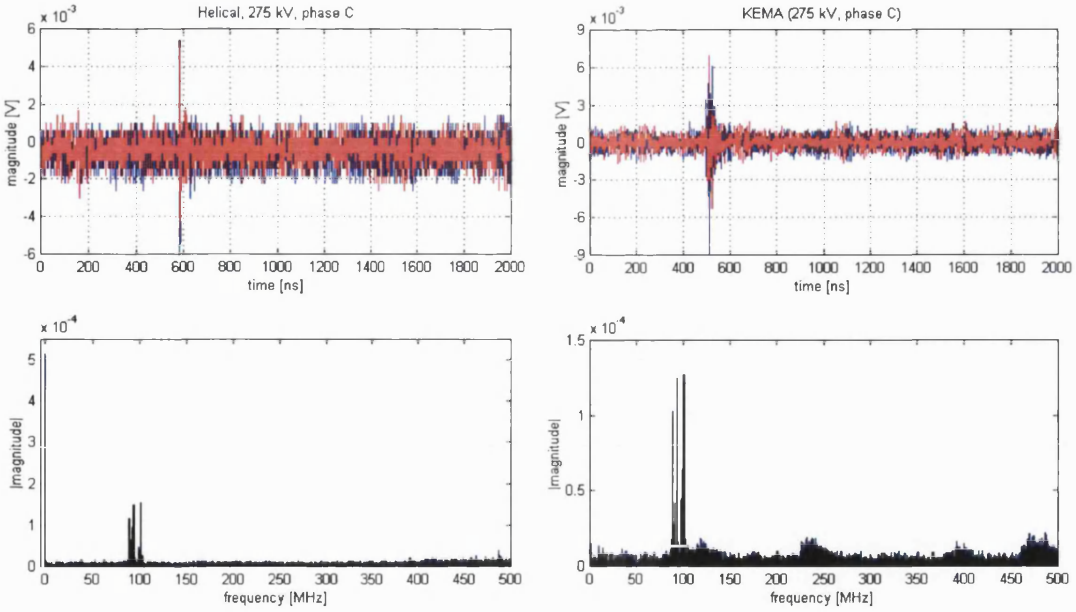


Figure 7.14: 3rd group of signals measured by the helical antenna H1 and the KEMA probe; triggered on helical antenna signals

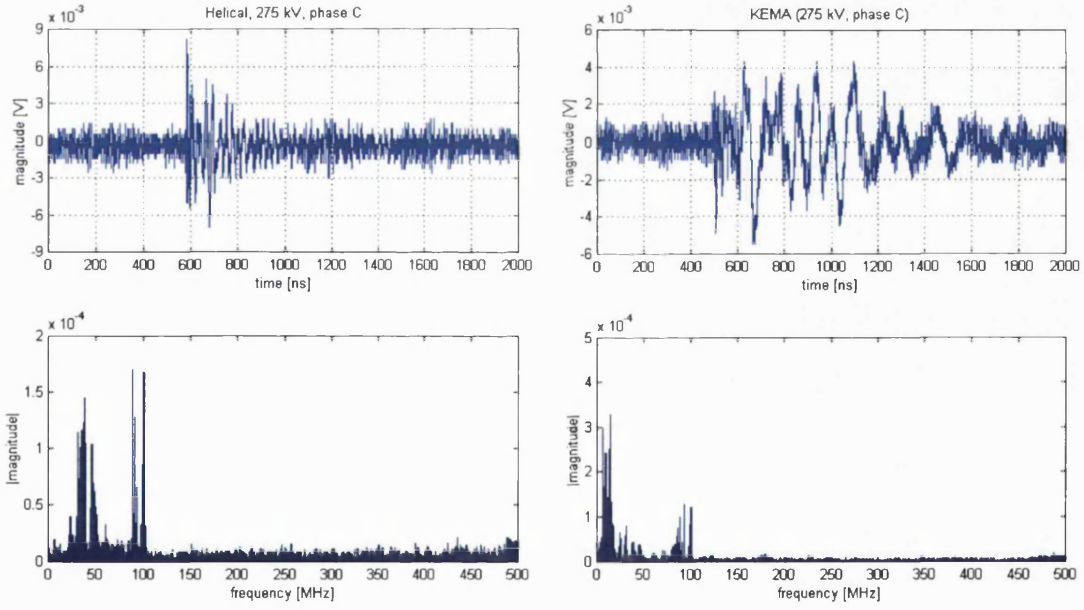


Figure 7.15: 4th group of signals measured by the helical antenna H1 and the KEMA probe; triggered on helical antenna signals

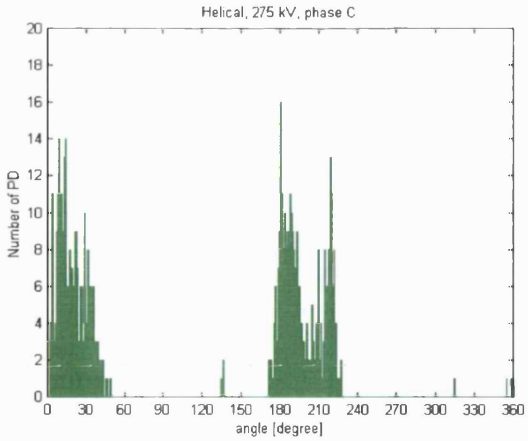


Figure 7.16: Distribution of the number of pulses regarding the voltage phase angle for H1

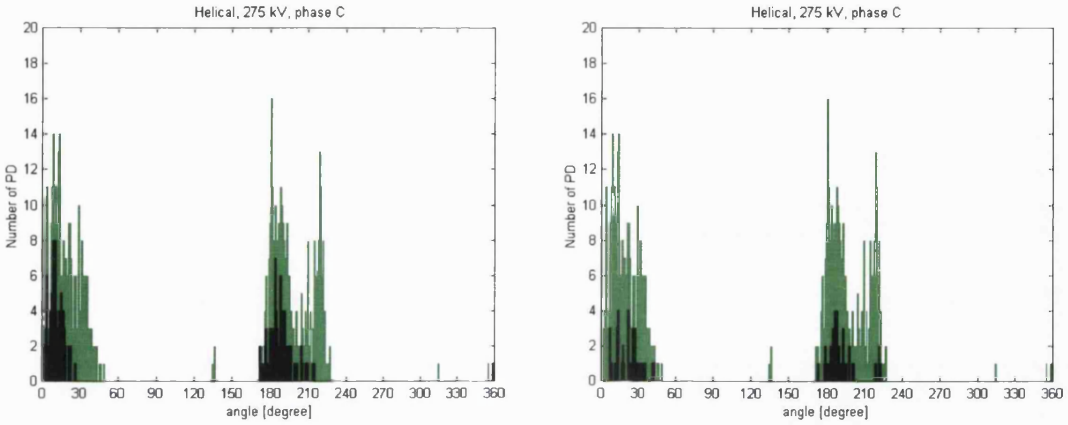


Figure 7.17: Distribution of the number of pulses regarding the voltage phase angle for the 1st and the 2nd group of signals

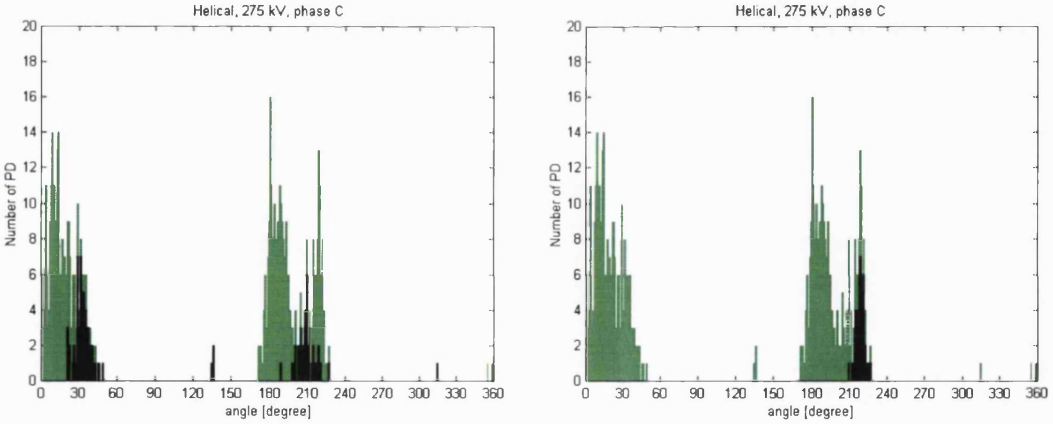


Figure 7.18: Distribution of the number of pulses regarding the voltage phase angle for the 3rd and the 4th group of signals

7.4. CONCLUSIONS ON WIDE BAND MEASUREMENTS

Based on presented results, the following conclusions can be drawn:

- Yagi antenna enables detection of multiple discharges. Correlation with the KEMA signals allows one type to be discarded. The dominant spectral content of all captured signals is below 100 MHz.
- Helical antenna also enables detection of multiple discharges. Pulses with dominant spectral content below 100 MHz are similar to pulses obtained by Yagi antenna. Pulses with dominant spectral content above 200 MHz, indicative of PDs in oil, were detected only in phase C on LV and HV side of the transformer.

Since only the helical antenna detected pulses with dominant frequencies above 200 MHz, measurements obtained by it will be considered in the further analysis. More than ten thousand measurements were captured in the substation and each measurement consists of 5000 samples, and hence it is appropriate to find a method that enables extraction of useful information from signals. The following chapter deals with feature extraction and group formation problem.

8. DATA MINING

8.1. INTRODUCTION

Data mining implies data analysis and discovery algorithms to find useful information from vast amounts of data. In many applications, data, which is the subject of analysis and processing in data mining, is multidimensional, and presented by a number of features. The so-called "curse of dimensionality" pertinent to many learning algorithms, denotes the drastic increase of computational complexity and classification error with data having a large amount of dimensions. Hence, an attempt is often made to reduce the dimensionality of the feature space is often tried to be reduced before classification is undertaken.

In practice, the two primary goals of data mining tend to be prediction and description. Prediction involves using some variables or fields in a data set to predict unknown or future values of other variables of interest. Description, on the other hand, focuses on finding patterns describing the data that can be interpreted by humans. The goals of prediction and description are achieved by using the following primary data-mining tasks [85]:

1. *Classification*, in which a function is learned which maps or classifies a data item into one of several predefined classes.
2. *Regression*, in which a function is learned which maps a real-valued prediction variable.
3. *Clustering*, in which a common descriptive task is used to identify a finite set of categories or clusters to describe the data.
4. *Summarisation*, in which a compact description is found for a subset of data.
5. *Dependency modelling*, in which a model is found which describes significant dependency between variables.
6. *Change and deviation detection*, in which significant changes in data set are discovered from previously measured values.

Considering how different authors describe data mining, it is evident that one is far from a universal agreement on the definition of data mining or even what constitutes data mining. Is data mining a form of statistics enriched with learning theory or it is a revolutionary new concept? In someone's view, most data-mining problems and corresponding solutions have roots in classical data analysis. Data mining has its origins in various disciplines, of which the

two most important are statistics and machine learning. Statistics has its roots in mathematics, and therefore, a desire to establish that something is sensible on theoretical grounds before testing it in practice. In contrast, the machine-learning community has its origins very much in computer practice. This has led to a practical orientation, a willingness to test something out to see how well it performs, without waiting for a formal proof of effectiveness.

8.2. DATA MINING PROCESS

By definition, data mining is the nontrivial process of extracting valid, previously unknown, comprehensible, and useful information from large databases and its application. It is an exploratory data analysis, trying to discover useful patterns in data that are not obvious to the data user.

In general, the data mining process iterates through five basic steps [86]:

1. **Data selection**, which consists of choosing the goal and the tools of the data mining process, identifying the data to be mined, then choosing appropriate input attributes and output information to represent the task.
2. **Data transformation operations**, which include organizing data in desired ways, converting one type of data to another (e.g., from symbolic to numerical) defining new attributes, reducing the dimensionality of the data, removing noise, "outliers", normalizing, if appropriate, deciding strategies for handling missing data.
3. **Data mining** step per se, in which the transformed data is subsequently mined, using one or more techniques to extract patterns of interest (The user can significantly aid the data mining method by correctly performing the preceding steps.).
4. **Result interpretation and validation**, in which the data mining application tests its robustness, using established estimation methods and unseen data from the database. This is for understanding the meaning of the synthesized knowledge and its range of validity. The extracted information is also assessed more subjectively, by comparing it with prior expertise in the application domain.
5. **Incorporation of the discovered knowledge**, which consists of presenting the results to the decision maker, who may check/resolve potential conflicts with previously believed or extracted knowledge and apply the new discovered patterns.

Figure 8.1 presents the whole process schematically by showing what is happening with the data. It is preprocessed, mined and postprocessed, the result being a refinement in the knowledge about the application. The data mining process is iterative, interactive, and very much a trial-and error activity.

Visualization plays an important role. Because someone finds it difficult to emulate human intuition and decision-making on a machine, the idea is to transform the derived knowledge into a format that is easy for humans to digest, such as images or graphs. Then, someone relies on the speed and capability of the human user visual system to spot what is interesting, at every step of the data mining process i.e., preliminary representation of data, domain specific visualization, or result presentation.

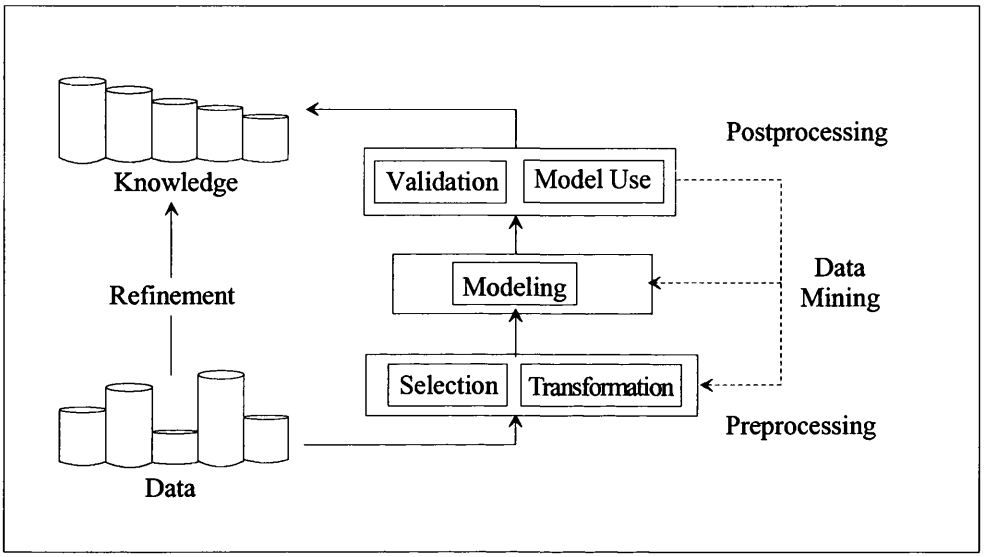


Figure 8.1: Data mining process

8.3. DATA MINING TASKS

Depending mainly on the application domain and on the interest of the miner, one can identify several types of data mining tasks for which data mining offers possible answers [86]. They are presented in the following Sections assuming that they are usually implicit in the process.

8.4. SUMMARIZATION

This task aims to produce compact and characteristic descriptions for a given set of data. It can take multiple forms: numerical (simple descriptive statistical measures like means, standard deviations), graphical forms (histograms, scatter plots), or the form of if-then rules. It may provide descriptions about objects in the whole database or in selected subsets of it.

8.4.1. Clustering

A clustering problem is an unsupervised learning problem that aims at finding clusters of similar objects sharing a number of interesting properties. It may be used in data mining to evaluate similarities among data, to build a set of representative prototypes, to analyze correlations between attributes, or to automatically represent a data set by a small number of regions, preserving the topological properties of the original input space.

8.4.2. Classification

A classification problem is a supervised learning problem in which the output information is a discrete classification i.e., given an object and its input attributes, the classification output is one of a possible mutually exclusive classes of the problem. The aim of the classification task is to discover some kind of relationship between the input attributes and the output class, so that the discovered knowledge can be used to predict the class of a new unknown object.

8.4.3. Regression

A regression problem is a supervised learning problem of building a more or less transparent model in which the output information is a continuous numerical value or a vector of such values rather than a discrete class. Then, given an object, it is possible to predict one of its attributes by means of the other attributes by using the built model. The prediction of numeric values may be done by classical or more advanced statistical methods and by symbolic methods often used in the classification task.

8.4.4. Dependency Modelling

A dependency modelling problem focuses on discovering a model that describes significant dependencies among attributes. These dependencies are usually expressed as if-then rules in the form "if antecedent is true then consequent is true," where both the antecedent and the consequent of the rule may be any combination of attributes, rather than having the same output in the consequent, similar to the case of the classification rules.

8.4.5. Deviation Detection

This is the task focusing on discovering the most significant changes or deviations in the data between the actual content of the data and its expected content (previously measured) or normative values. It includes searching for temporal deviations (important changes in data with time) and searching for group deviations (unexpected differences between two subsets of data).

8.4.6. Temporal Problems

In certain applications, it is useful to produce rules that explicitly take into account the role of time. There are databases containing temporal information that may be exploited by searching for similar temporal patterns in data or learn to anticipate some abnormal situations in data.

8.4.7. Causation Modelling

This is a problem of discovering relationships of cause and effect among attributes. A causal rule of type if-then indicates not only that there is a correlation between the antecedent and the consequent of the rule, but also that the antecedent causes the consequent.

8.5. DATA MINING TECHNIQUES

The enumerated types of data mining tasks are based on a set of important techniques originating in artificial intelligence paradigms, statistics, information theory, machine learning, reasoning with uncertainty (fuzzy sets), pattern recognition, or visualization [86]. Thus, a data mining software package is supported to varying degrees by a set of technologies, which nearly always include the following.

8.5.1. Tree and Rule Induction

Machine learning (ML) is the centre of the data mining concept due to its ability to gain physical insight into a problem, and participates directly in data selection and model search steps. The ML field is basically concerned with the automatic design of if-then rules similar to those used by human experts. Decision tree induction: the best known ML framework was found to be able to handle large-scale problems due to its computational efficiency, to provide interpretable results, and, in particular, able to identify the most representative attributes for a given task.

8.5.2. Association Rules

Association rule generators are a powerful data mining technique used to search through an entire data set for rules revealing the nature and frequency of relationships or associations between data entities. The resulting associations can be used to filter the information for human analysis and possibly to define a prediction model based on observed behaviour.

8.5.3. Clustering Methods

These are often used in the data selection preprocessing step due to their property of learning unsupervised similarities between objects and reducing the search space to a set of most important attributes for the application, or to a finite set of objects "alike." The most frequently used clustering method is the k-means method, which identifies a certain number of groups of similar objects; it may be used in combination with the nearest-neighbour rule, which classifies any new object in the group most similar (most near) to it. This method may also be used in order to identify outliers in a database.

8.5.4. Artificial Neural Networks

ANN's are recognized in the automatic learning framework as universal approximators, with massively parallel computing characteristics and good generalization capabilities, but also as black boxes due to the difficulty in obtaining an insight into the relationship learnt.

They are used within the data mining step to:

- Generate a regression model that can predict future behaviour on the basis of a database with input/output pairs of continuous numerical historical information (the neural network acts like mapping, associating numerical outputs to any new object of known attributes values)
- Automatically represent a data set by a small number of representative prototypes, preserving the topological properties of the original attribute space (unsupervised learning).

8.5.5. Statistical Techniques

These include linear regression, discriminant analysis, or statistical summarization.

8.5.6. Visualization Techniques

These include histograms (estimate the probability distribution for a certain numerical attribute given a set of objects), scatter plots (provide information on the relation between two numerical attributes and a discrete one), three-dimensional maps, dendrograms (a correlation analyses between attributes or objects), etc.

8.5.7. Additional Tools

In addition, some DM packages include:

- Genetic algorithms (optimization techniques based on the concepts of genetic combination, mutation and natural selection).

- Sequential patterns discovery (group objects with the same succession of given attribute values over a time period).
- Time-series similarity (detect similar time-series over a period of time).
- Bayesian belief networks (graphical models that encode probabilistic relationships among variables of interest, systems able to learn causal relationships).
- Neuro-fuzzy systems (fuzzy inference systems that incorporate the learning and generalization abilities of neural networks).

8.5.8. Common Features

Even if one considers data mining tools to be like toolboxes of multiple techniques able to perform a complete data analysis, the reality is not yet so, with the market presently offering only partially equipped products.

DM techniques are different from one other in terms of problem representation, parameters to optimize, accuracy, complexity, run time, transparency, and interpretability. Common features for all the techniques include:

- Making a compromise between accuracy and complexity by means of pruning techniques.
- Enhancing the comprehensibility of derived patterns.
- Fighting to avoid the overfitting of a problem that appears when the model to be extracted is too complex with respect to the information provided in the learning set.

8.6. DM Operational Aspects

The success of mining some data is induced by a list of factors [86].

8.6.1. The Right Tools

A distinctive feature of data mining software is the quality of its algorithms, the effectiveness of the techniques, and sometimes their speed. In addition, the efficiency of using the hardware, the operating system, the database resources, and the parallel computing are having an influence on the process. Moreover, it turns out that the particular set of tools useful in a given application is highly dependent on the practical problem.

Thus, at the prototype stage, it is useful to have available a broad enough set of techniques so as to identify interesting applications. However, in the final product used for actual field implementation, it is often possible to use only a small subset of the latter tools.

Customizing data mining techniques to the application domain and using methods that are a reliable means to achieving the proposed goal may enhance the process of extracting useful information.

8.6.2. The Right Data

The data to be mined should contain information worth mining i.e., consistent, clean, and representative for the application. Of course, it is useless to apply data mining to an invalid database with high measurement or estimation data errors, or to try to precisely estimate numerical outputs that present, a priori, high noise. A data mining tool ideally explains as much information as is stored in the data that is mined (a derived model is strongly dependent on the learning set used), and sometimes it is not what is in the data that matters for an application (wrong attributes, wrong selected sample).

An important part of data mining errors in the results are due to uncertainties in modelling and generation of objects in certain databases in discord with the real probabilities of phenomena appearances in the system. This is why data mining errors often do not have a meaning by themselves; they just provide a practical means to compare efficiencies of different criteria applied to the same database.

8.6.3. The Right People

Regardless of what many producers of DM tools claim, data mining is not as yet an automatic operation with little or no human intervention. On the contrary, the human analyst plays an important role, mostly in the areas of data selection and data/knowledge interpretation. The data miner should have an understanding of the data under analysis and the domain or industry to which it pertains. It is more important for the mining process to embrace the problems the application is meant to solve rather than to incorporate in the data mining software the latest technologies.

8.6.4. The Right Application

Almost always, a problem well posed is already a partially solved problem. It is important to clearly define the goals and choose the appropriate objectives so as to yield a significant impact on the underlying decision making process.

8.7. DIMENSIONALITY REDUCTION

8.7.1. Introduction

The main theme for simplifying the data is dimension reduction, and the main question is whether some of these prepared and pre-processed data can be discarded without scarifying the quality of results. There is one additional question about the techniques for data reduction: Can the prepared data be reviewed and a subset found in a reasonable amount of time and space?

It is important to make a distinction between feature selection and feature extraction. The term feature selection refers to algorithms that select, hopefully, the best subset of the input feature set. Methods that create new features based on transformations or combinations of the original feature set are called feature extraction algorithms. However, the terms feature selection and feature extraction are used interchangeably in the literature. Note that often feature extraction precedes feature selection; first, features are extracted from the sensed data (e.g., using principal component or discriminant analysis) and then some of the extracted features with low discrimination ability are discarded. The choice between feature selection and feature extraction depends on the application domain and the specific training data which is available. In addition, the retained features may be important for understanding the physical process that generates the patterns. On the other hand, transformed features generated by feature extraction may provide a better discriminative ability than the best subset of given features, but these new features (a linear or a nonlinear combination of given features) may not have a clear physical meaning.

In many situations, it is useful to obtain a two- or three-dimensional projection of the given multivariate data (n x k pattern matrix) to permit a visual examination of the data. Several graphical techniques also exist for visually observing multivariate data, in which the objective is to exactly depict each pattern as a picture with k degrees of freedom, where k is the given number of features.

Since the work of this thesis is related to feature extraction, some of the commonly used methods for feature extraction are briefly discussed.

8.7.2. Feature Extraction

Feature extraction methods determine an appropriate subspace of dimensionality m (either in a linear or a nonlinear manner) in the original feature space of dimensionality k ($m \leq k$). Linear transforms, such as principal component analysis, factor analysis, linear discriminant analysis, and projection pursuit have been widely used in pattern recognition for feature extraction and dimensionality reduction. The best known linear feature extractor is the principal component analysis (PCA) or Karhunen-Loève expansion that computes the m largest eigenvectors of the k x k covariance matrix of the n x k -dimensional patterns [87]. Since PCA uses the most expressive features (eigenvectors with the largest eigenvalues), it effectively approximates the data by a linear subspace using the mean-square-error criterion.

Independent component analysis (ICA) has been successfully used for blind-source separation [88] which extracts linear feature combinations that define independent sources. This demixing is possible if at most one of the sources has a Gaussian distribution. PCA on the other hand, is an unsupervised linear feature extraction method, discriminant analysis uses the

category information associated with each pattern for (linearly) extracting the most discriminatory features. In discriminant analysis, interclass separation is emphasised by replacing the total covariance matrix in PCA by a general separability measure like the Fisher criterion, which results in finding the eigenvectors of $S_w^{-1} S_b$ (the product of the inverse of the within-class scatter matrix, S_w, and the between-class scatter matrix, S_b).

Multidimensional scaling (MDS) is a nonlinear feature extraction technique [87]. It aims to represent a multidimensional dataset in two or three dimensions such that the distance matrix in the original k-dimensional feature space is preserved as faithfully as possible in the projected space. Various stress functions are used for measuring the performance of this mapping [87]. A problem with MDS is that it does not give an explicit mapping function and hence it is not possible to place a new pattern in a map which has been computed for a given training set without repeating the mapping.

A feed-forward neural network offers an integrated procedure for feature extraction and classification; the output of each hidden layer may be interpreted as a set of new, often nonlinear, features presented to the output layer for classification. In this sense, multilayer networks serve as feature extractors [89].

The Self-Organizing Map (SOM), or Kohonen Map [90], can also be used for nonlinear feature extraction. In SOM, neurons are arranged in a k-dimensional grid, where k is usually 1, 2, or 3. Each neuron is connected to all the k input units. The weights on the connections for each neuron form a k-dimensional weight vector. During training, patterns are presented to the network in a random order. At each presentation, the winner whose weight vector is the closest to the input vector is first identified. Then, all the neurons in the neighbourhood (defined on the grid) of the winner are updated such that their weight vectors move towards the input vector. Consequently, after training is done, the weight vectors of neighbouring neurons in the grid are likely to represent input patterns which are close in the original feature space. Thus, a "topology-preserving" map is formed. When the grid is plotted in the original space, the grid connections are more or less stressed according to the density of the training data. Thus, SOM offers a k-dimensional map with a spatial connectivity, which can be interpreted as feature extraction.

In the following, the PCA and method of moments are briefly described since they are used as dimensionality reduction techniques in this thesis.

8.7.3. Principal Component Analysis

8.7.3.1. Introduction

The most popular statistical method for dimensionality reduction of a large data set is the Karhunen-Loève (K-L) method, also called Principal Component Analysis (PCA) [85]. PCA is

a method of transforming the initial data set represented by vector samples with derived dimensions. The goal of this transformation is to concentrate the information about the differences between samples into a small number of dimensions. The basic idea can be described as follows: A set of n-dimensional vector samples $X=\{x_1, x_2, x_3, \ldots, x_m\}$ should be transformed into another set $Y=\{y_1, y_2, \ldots, y_m\}$ of the same dimensionality, but y-s have the property that most of their information content is stored in the first few dimensions. This will allow one to reduce the data set into a smaller number of dimensions with low information loss.

8.7.3.2. Principal Components

The transformation is based on the assumption that high information corresponds to high variance. Thus, if one wants to reduce a set of input dimensions **X** to a single dimension **Y**, one should transform **X** into **Y** as a matrix computation:

$$\mathbf{Y} = \mathbf{A} \times \mathbf{X} \tag{8.1}$$

A is chosen such that **Y** has the largest variance possible for a given set. The single dimension **Y** obtained in this transformation is called the first principal component which is an axis in the direction of maximum variance. It minimizes the distance of the sum of squares between data points and their projections on the component axis.

In practice, it is not possible to determine matrix **A** directly and hence the computation of the covariance matrix **S** is a first step in features transformation. Matrix **S** is defined as

$$\mathbf{S}_{nxn} = \frac{1}{(n-1)}\left[\sum_{j=1}^{n}\left(x_j - \bar{x}\right)^T\left(x_j - \bar{x}\right)\right]$$

$$\text{where} \quad \bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j \tag{8.2}$$

The eigenvalues ($\lambda_1, \lambda_2, \ldots\lambda_n$) of the covariance matrix **S** for the given data should be calculated in the next step, where:

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0. \tag{8.3}$$

The eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots\mathbf{e}_n$ correspond to eigenvalues $\lambda_1, \lambda_2, \ldots\lambda_n$, and they are called the principal axes.

Finally, the m eigenvectors ($\mathbf{e}_1, \mathbf{e}_2, \ldots\mathbf{e}_m$) corresponding to the m largest eigenvalues of **S** define a linear transformation from n-dimensional space to an m-dimensional space in which the features are uncorrelated.

Principal axes are new, transformed axes of n-dimensional space, where the new variables are uncorrelated, and variance of the i-th component is equal to i-th eigenvalue. Because λ_i's are

sorted, most of the information about the data set is concentrated in a few first principal components.

The fundamental question that needs to be asked is how many of the principal components are needed to get a good representation of the data? In other words, what is the effective dimensionality of the data set? There is no definitive answer to this question. The easiest way to answer the question is to analyze the proportion of variance. Dividing the sum of the first m eigenvalues by the sum of all variances (all eigenvalues), will get the measure for the quality of representation based on the first m principal components. The result is expressed as a percentage, and if, e.g., the projection accounts for over 80% of the total variance, it is considered to be good. More formally, this ratio can be expressed in the following way.

The criterion for features selection is based on the ratio of the sum of the m largest eigenvaluse of **S** to the trace of **S**. This is a fraction of the variance retained in the m-dimensional space. If the eigenvalues are labelled so that $\lambda_1 \geq \lambda_2, \geq \ldots \geq \lambda_n$, then the ratio can be written as:

$$R = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{8.4}$$

When the ratio R is sufficiently large (greater than the threshold value), all analyses of the subset of m features represent a good initial estimate of n-dimensionality space.

Another common approach uses a scree[1] graph, in which eigenvalues are ordered from the largest to the smallest; a scree plot is a plot of λ_i versus i – the magnitude of an eigenvalue versus its number. The number of components m is chosen to be the point at which the line in the scree graph is "steep" to the left but "not steep" to the right. The number of components is taken to be the point at which the remaining eigenvalues are relatively small and all about the same size. The main problem with these approaches is that they are very subjective.

8.7.3.3. Simple example illustrating PCA

A simple example can be used to illustrate the principle behind the PCA. In the example, the number of characters will be reduced from two to one. In real applications, the method is used to reduce the number of characters from many to two or three.

In Figure 8.2, a number of objects characterised by only two characters is plotted. Through these points, a line needs to be drawn so that the variance among the points when projected onto this line will be as large as possible (this line is called the first principal component). This ensures that as much information as possible about the original data set will be retained. When

[1] Scree is the rock debris at the bottom of a cliff [87].

this line has been found, all the points are projected onto it. On this line (i.e. the reduced character space) it may be possible to distinguish clusters by visual inspection. This new line, or character, can be interpreted in terms of the contributions that the original characters have made to it.

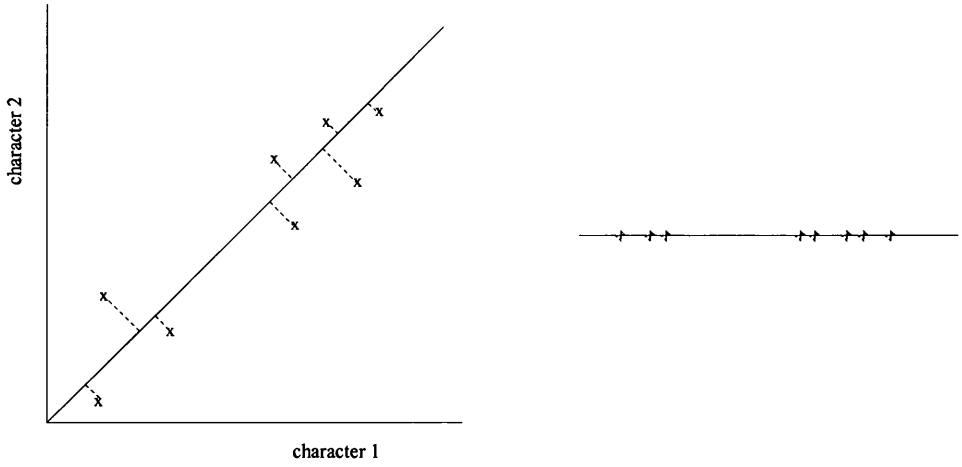


Figure 8.2: An example illustrating PCA

When PCA is used to reduce the number of characters from many to two or three, not only does the first but also the second and third principal components are calculated, and the points are projected not onto a line but onto a two- or three-dimensional character space.

8.7.4. Method of moments

One is often concerned with the problem of characterization of a signal by associating with its most significant parameters (e.g. bandwidth, energy content, peak value, duration, etc.) rather than obtaining a complete specification of its character. Thus, one seeks appropriate mappings from the signal space into numerical values. The primary motivation for extracting certain features from signals is the desire to ascribe performance measures by which various systems can be described.

Amplitude and polarity independent are suggested by the need to separate PD pulses on the basis of their shape, independently on polarity and amplitude. Amplitude independent needs normalization of PD pulses before computing PD features. Thus, if $s(t)$ is the recorded signal, normalization can be achieved by:

$$\tilde{s} = \frac{s}{\|s\|} = \frac{s}{\sqrt{\sum_{i=1}^{N} s(t_i)^2}}, \tag{8.5}$$

This is the reciprocal of the Euclidian norm of the vector containing PD pulse samples. Polarity independent can be satisfied by resorting to a quadratic functional applied to the pulse shape.

The concept of measuring the width of a probability density function in terms of its second central moment as in the probability theory [91] is a useful tool in the analysis of signals. Unlike a probability density function a signal $s(t)$ is not necessarily non-negative and the second moment of $s(t)$ may be meaningless for establishing the width of a signal having a high oscillatory behaviour. To avoid this problem, the second central moment of $s^2(t)$ is considered and since the zeroth moment of $s^2(t)$ is not necessarily unity, the second moment is normalized and gives a measure of signal duration σ_T:

$$\sigma_T = \sqrt{\sum_{i=1}^{N} (t_i - t_0)^2 \cdot \tilde{s}_i(t_i)^2} \tag{8.6}$$

where

$$t_0 = \sum_{i=1}^{N} t_i \cdot \tilde{s}_i(t_i)^2 \tag{8.7}$$

is the normalized first moment of $s^2(t)$.

Another analogy is apparent when the second moment is considered as the moment of inertia of a mass distribution given by $s^2(t)$. For this reason, σ_{Ti} is often referred to as the "radius of gravitation" of the signal $s(t)$. Parameter t_0 is the "gravity centre" of the normalized signal. Similarly, the bandwidth occupied by a signal in terms of the second central moment of the square of the magnitude of its Fourier transform can be measured as:

$$\sigma_F = \sqrt{\sum_{i=1}^{N} f_i^2 \cdot \left|\tilde{S}_i(f_i)\right|^2} \tag{8.8}$$

where $f_0 = 0$ because of the symmetry of $S(f)$ about the origin for real signals.

Two real numbers obtained by equations (8.6) and (8.8) localize the signal in the time-frequency plane. Therefore, signals represented with σ_T and σ_F can be given as dots in the Cartesian plane. PD-pulse signals having similar shape can be synthesized in well-defined areas of σ_T, σ_F plane.

8.8. SELF-ORGANISING MAP

The Self-Organizing Map (SOM) [92] developed by professor Kohonen, is one of the most popular neural network models. The SOM algorithm is based on unsupervised competitive learning, which means that the training is entirely data-driven and that the neurons of the map compete with each other. Supervised algorithms, like the Multi-Layered Perceptron (MLP), require that the target values for each data vector are known, but the SOM does not have this limitation.

The SOM can be applied for "unearthing" the hidden structure or inherent characteristics of a highly complex or multi-dimensional data and displaying it in a discernible and comprehensible manner. Due to the foregoing capability of SOM, it has been applied successfully in various fields [92] such as knowledge extraction, classification, monitoring and modelling of complex processes, analysis of pulp and paper industries and correlation hunting.

8.8.1. The structure

The SOM is formed by neurons located on a regular, usually one- or two-dimensional grid. A higher dimensional grid can also be utilised, but it is better to avoid it due to the difficulty in visualisation. Besides the input layer, the Kohonen networks contain the Kohonen layer, as shown in Figure 8.3. The SOM essentially defines a mapping from the input space, $\Re^n$, on to a two-dimensional array of neurons. Every neuron, i, is associated with an n-dimensional reference vector, $\mathbf{m}_i = [w_{i1}, w_{i2}, w_{i3}, \ldots, w_{in}]^T \in \Re^n$; where n is the dimension of the input space (i.e. number of components within an input vector). Hence, an input vector, $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_n]^T \in \Re^n$, is connected to all neurons in parallel via scalar values w_{ij}, $j = 1, 2, \ldots, n$, which are different for every neuron.

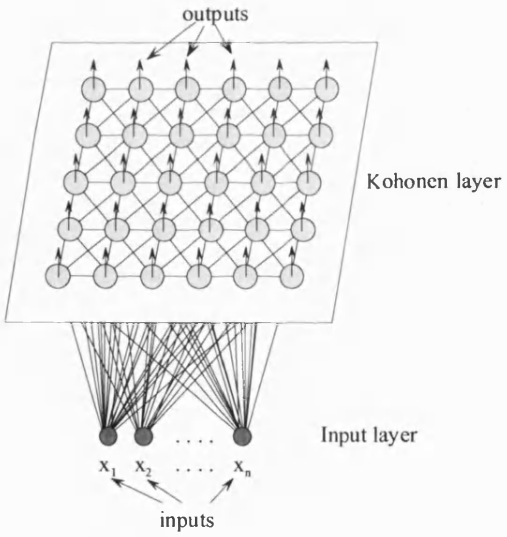


Figure 8.3: Two-dimensional SOM with neurons arranged in a hexagonal lattice

In addition, each neuron is associated with adjacent neurons of the SOM by a neighbourhood relation, which dictates the "area of influence" in the grid. Therefore, the immediate neighbours of neuron i, which are closest to neuron i, are associated with the first-neighbourhood, N_{i1}, of neuron i. The neighbourhood relations of different sizes in rectangular and hexagonal lattices are illustrated in Figure 8.4 [93].

The training process of SOM comprises of two stages, as explained in the following sections.

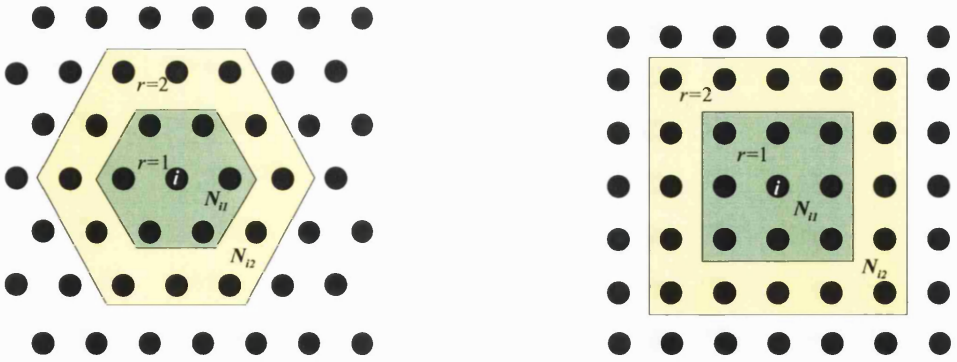


Figure 8.4: The neighbourhood relation of hexagonal and rectangular grid

8.8.1.1. Initialisation

In the basic SOM algorithm, the topological relations and the number of neurons are fixed at the outset. Typically, one of the three following initialization procedures is used:

1. Random initialization, where the reference vectors are initialised using arbitrary values; all initially unordered reference vectors will become organised after a few hundred training time-steps.
2. Sample initialization, where the weight vectors are initialized with random samples drawn from the input data set. This is done so as to have a rough approximation of the density function of the input space before the start of the training process.
3. Linear initialization, where the weight vectors are initialized in an orderly fashion along the linear subspace spanned by the two principal eigenvectors of the input data set.

8.8.1.2. The training process

Once reference vectors of SOM have been initialised, the training of SOM commences by first by choosing an input vector, $\mathbf{x}_k$, randomly from the input data. Comparison is then performed with the computation of similarity measures, between $\mathbf{x}_k$, and all reference vectors of SOM. The similarity is usually defined by a distance measure, usually Euclidean distance:

$$d(\mathbf{x}_k, \mathbf{m}_j) = \sqrt{\sum_n (\mathbf{x}_{kn} - w_{jn})^2} \tag{8.9}$$

The best match of $\mathbf{x}_k$ on the SOM grid is known as the best matching unit (BMU), denoted as w_{ij}, and the best matching vector is with $\mathbf{m}_c$. The neuron c is hence the location of the "response" of SOM towards the input vector, $\mathbf{x}_k$. In simple terms, the input vector, $\mathbf{x}_k$, is "mapped" onto the SOM at location c. Formally the BMU is defined as the neuron for which:

$$d = \|\mathbf{x} - \mathbf{m}_c\| = \min_i \{ \|\mathbf{x} - \mathbf{m}_i\| \} \tag{8.10}$$

where $\|\cdot\|$ is the distance measure.

After finding the BMU, the weight vectors of the SOM are updated. The weight vectors of the BMU and its topological neighbours are moved closer to the input vector in the input space. This adaptation procedure stretches the BMU and its topological neighbours towards the sample vector. This is illustrated in Figure 8.5 where the input vector given to the network is marked by an x.

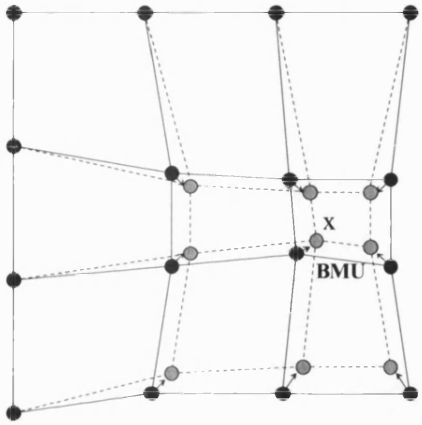


Figure 8.5: Influence of input vector towards its BMU and neighbours

The SOM update rule for the weight vector of the unit i is:

$$\mathbf{m}_i(t+1) = \begin{cases} \mathbf{m}_i(t) + h_{ci}(t)\left[\mathbf{x}(t) - \mathbf{m}_i(t)\right]; & i \in N_C(t) \\ \mathbf{m}_i(t); & i \notin N_C(t) \end{cases} \tag{8.11}$$

where t denotes time. The $x(t)$ is the input vector randomly drawn from the input data set at time t and $h_{ci}(t)$ the neighbourhood kernel around the winner unit c at time t. The neighbourhood kernel is a non-increasing function of time and of the distance of unit i from the winner unit c. It defines the region of influence that the input sample has on the SOM. The kernel comprises of two parts, the neighbourhood function $h(d,t)$ and the learning rate function $\alpha(t)$:

$$h_{ci}(t) = \alpha(t) h_{ci}\left(\|r_c - r_i\|, t\right) \tag{8.12}$$

where r_i is the location of unit i on the map grid. The learning rate function lies in the interval between 0 and 1. For convergence, it is necessary that $h_{ci}(t) \to 0$ when $t \to \infty$, and, with increasing $\|r_c - r_i\|$, $h_{ci} = 0$.

There are two common choices for the neighbourhood function: rectangular function and Gaussian function, as illustrated in Figure 8.6 [93].

For rectangular function, the neighbourhood kernel can be defined as:

$$h_{ci}(t) = \begin{cases} \alpha(t), & i \in N_c(t) \\ 0, & i \notin N_c(t) \end{cases} \tag{8.13}$$

For Gaussian function, the neighbourhood kernel can be defined as:

$$h_{ci}(t) = \alpha(t) \cdot \exp\left(-\frac{\|r_c - r_i\|^2}{2\sigma^2(t)} \right) \tag{8.14}$$

Usually the neighbourhood radius is larger than the unity in the beginning and is gradually decreased (linearly) to unity during the training.

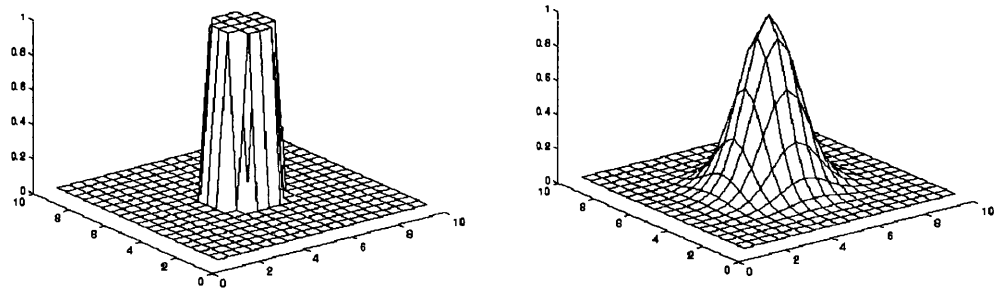


Figure 8.6: Neighbourhood functions of SOM (left rectangular and right Gaussian function)

In summary, the training process of SOM comprises of the following steps:

Step 1: Initialise reference vectors of all neurons at time-step, $t = 0$.

Step 2: At each time-step, t, select one input vector, $x(t)$, randomly from the input data.

Step 3: At each time-step, t, determine the BMU, m_c, for input vector, $x(t)$ according to Equation (8.10)

Step 4: Update all reference vectors according to Equation (8.11).

Step 5: Check if user-defined iterations and criteria have been reached. If YES, then stop training; if NO, increase time-step, t, by 1 and repeat the process from step 2.

Various issues have to be considered for practical application of SOM; these will be discussed in Section 8.8.1.4.

8.8.1.3. Visualization

Because the weight vectors are ordered on grid, the SOM can be easily used for visualization. There are a variety of different kinds of visualization techniques available for the SOM.

An excellent approach of visualisation has been proposed by Ultsch [90], [94]-[95], which is known as the "unified-matrix" (u-matrix) method. The u-matrix method is essentially based on the Euclidean distances between SOM neurons that are located on the SOM grid. Essentially, it allows the visualisation of multi-dimensional reference vectors in a two-dimensional format, which is discernible and comprehensible to the end user. It effectively allows the visualisation of the hidden "structure" or inherent characteristics of the input data, which have been "learned"

or "mapped" by the trained SOM. Fundamental principles of the u-matrix method are presented in Section 8.8.3.

8.8.1.4. The configuration of SOM

The number of neurons:

A two-dimensional SOM in a rectangular shape is generally adequate for most problems. In addition, the neurons are arranged in hexagonal configuration for ease of visualisation. The total number of neurons required can be determined empirically via the following formula [96]:

$$M = 5 \times \sqrt{n} \tag{8.15}$$

where, M is the total number of neurons and n is the total number of input vectors. The number of neurons in the x- and y-dimension of the SOM (i.e. map-sizes) can be determined by calculating the two principal eigenvalues ($\beta_1$in β_2) of the covariance matrix of input data:

$$\frac{x}{y} = \sqrt{\frac{\beta_1}{\beta_2}} \tag{8.16}$$

However, it has been found [96] that (4 x *M)* or (9 x *M)*, is generally more appropriate for better visualisation of the training outcome. Consequently, the final number of neurons and dimensions of neural net is obtained by the following equation:

$$\begin{gathered} 2x \times 2y = 4 \times M \\ or \\ 3x \times 3y = 9 \times M \end{gathered} \tag{8.17}$$

The selection of optimum SOM

Given a SOM with a pre-determined configuration and pre-defined training parameters, how does one identify the number of training iterations needed in order to produce an optimum SOM for a set of input data? The answer is not so straight-forward due to the fact that no "targets" are available for the validation of SOM during the training process, similar to other types of unsupervised NNs. Therefore, a combination of mathematical and heuristic approaches must be adopted so as to identify the optimum SOM for a given set of input data in a systematic and unambiguous manner.

There is no single parameter which seamlessly measures the effectiveness of SOM for meeting the foregoing objectives. However, two SOM quality measures are available to calculate the ability of SOM for vector quantisation and generalisation, respectively [96]. The ability of SOM to approximate a given set of multi-dimensional input data is measured by the "average quantisation error" (ε_{AQ}). It is defined as the average Euclidean distance between input vectors and their BMUs. The ε_{AQ} will tend to decrease with an increase in the training iterations, since a longer training interval may lead to a gradual over-fitting of the training data.

The ability of SOM to generalise and preserve its topology is measured by the "topographic error" (ε_T). It is defined as the proportion of all input vectors for which their first and second BMUs are not adjacent to each other. The ε_T will tend to rise and fall at several stages throughout the training process, which are observed to correspond to the clarity of the hidden "structure" of the input data. However, a low ε_T does not necessarily mean an optimum SOM since the corresponding ε_{AQ} might still be quite high.

A group of identically configured SOMs are trained using the same training parameters but only differ in the number of training iterations, which are increased gradually by a fixed amount. The training process is restarted for every SOM, i.e. after the initialisation of reference vectors for the "rough ordering" phase and after the "rough ordering" phase for the "fine tuning" phase. Both ε_{AQ} and ε_T are calculated after the training is completed for every SOM; a likely scenario for the magnitude of these measures is illustrated in Figure 8.7. Several suitable candidates for the optimum SOM are selected from several "valleys" of the ε_T curve, as pointed out in Figure 8.7. Finally, the optimum SOM is selected from these candidates based on human visual inspection.

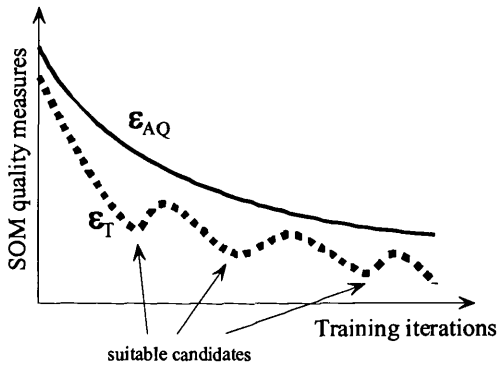


Figure 8.7: Selection of optimum SOMs based on several suitable candidates

8.8.2. Software Tools of SOM

Various software tools are available for those who are interested in using the SOM for data analysis. Among the SOM software tools are SOM_PAK [97] and its MATLAB implementation, and the SOM Toolbox [98], which are developed by the Neural Network Research Centre (NNRC) at the Helsinki University of Technology, Finland. Since the NNRC is established by Prof. Teuvo Kohonen, the founder of the SOM algorithm, the above-mentioned software tools are the most faithful implementation of the SOM theories. Hence, these software tools have been employed herein for conducting the research of data mining on partial discharge data, as reported in subsequent chapters of this thesis.

8.8.3. An Example of SOM application

The application of SOM for a simple data analysis is demonstrated. The input data contains 300 input vectors; there are 3 components in each input vector. The input data is generated by mixing three subsets, with each subset being constructed from 100 randomly generated vectors that are centred at [0 0 0], [3 3 3] and [9 0 0], respectively as shown in Figure 8.8.

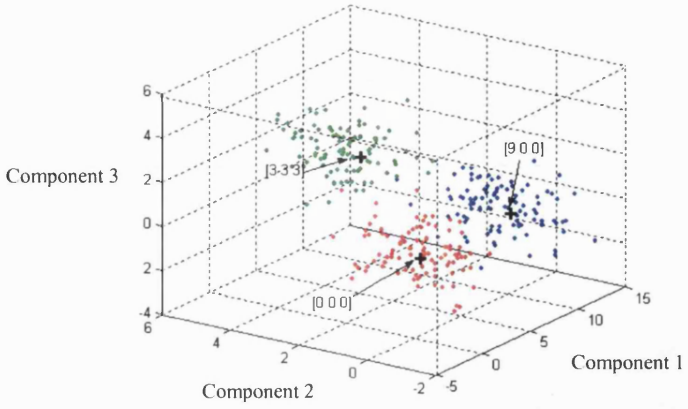
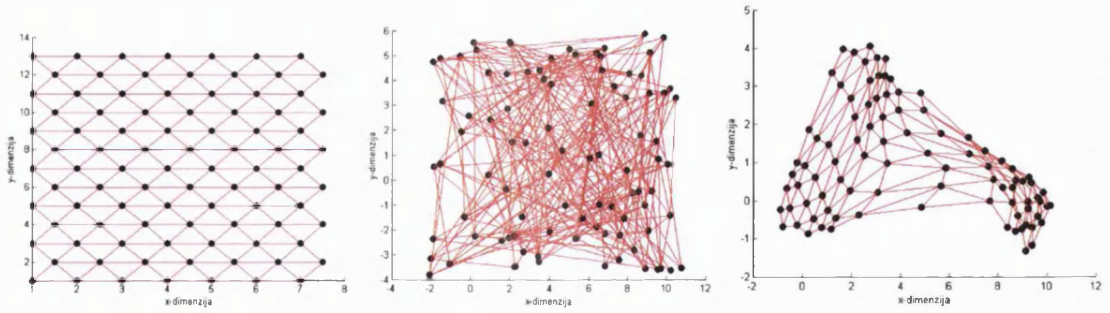


Figure 8.8: The input data where input vectors consists of three components

A SOM is configured in a hexagonal lattice of 7-by-13 neurons, as illustrated in Figure 8.9.a. Before the commencement of training, reference vectors of SOM are randomly initialised, resulting in the initial arrangement of SOM neurons being still "messy", as illustrated in Figure 8.9.b. Figure 8.9.c gives the situation of the arrangement of SOM neurons after the training phase.



a) A SOM consists of 24-by-14 neurons arranged in hexagonal lattice	b) Arrangement of SOM neurons after the random initialisation	c) Arrangement of SOM neurons after the training

Figure 8.9: Arrangement of neurons

Another approach of visualising the trained SOM is the visualisation of u-matrix and component planes, as illustrated in Figure 8.10. Colour bars on the right-hand-side of the u-matrix and component planes represent the magnitude, in which the colour bar on the u-matrix represents Euclidean distances between neurons and the colour bars on component planes represent the magnitude of each component of the reference vectors.

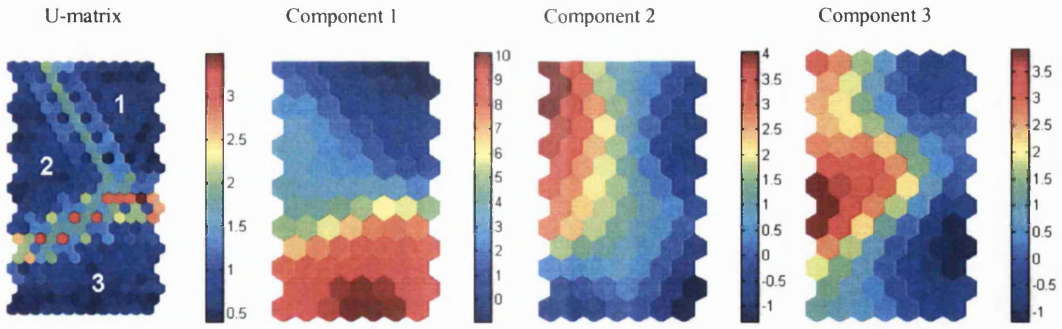


Figure 8.10: The u-matrix and component-planes

As observed from the u-matrix, the map is divided into three regions of smaller Euclidean distances, in which each region actually corresponds to a subset of the input data. High value of the Component 1 and low values of Components 2 and 3 indicate the third region in u-matrix. It corresponds to the input data with centre at [9 0 0]. Medium value of the Component 1 and high values of Components 2 and 3 correspond to the second region in u-matrix and input data with centre at [3 3 3]. Low values of all three components corresponds to the region 1 in u-matrix and the input data with centre at [0 0 0]. Figure 8.11 gives u-matrix in grey scale and the input data transform into u-matrix. Red hexagons presents input data with centre at [0 0 0], green hexagons input data with centre at [3 3 3], and blue hexagons input data with centre at [9 0 0].

The presented example shows that the hidden "structure" or inherent characteristics of data can be observed through the use of the u-matrix visualisation.

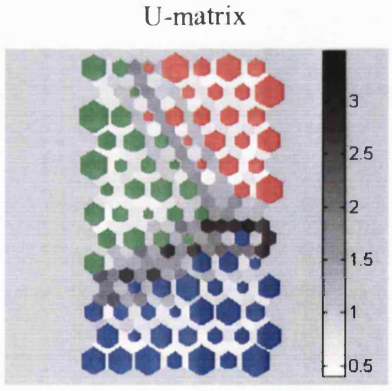


Figure 8.11: Input data "transformed" into u-matrix

8.9. CLUSTER ANALYSIS [99]- [101]

8.9.1. Introduction

The objective of cluster analysis (CA) is to group data into clusters so that the degree of association is strong between members of the same cluster and weak between members of different clusters. Thus each cluster describes the class to which its members belong. As a result, CA can reveal similarities in data, which may have been otherwise impossible to find. Examples of such applications would be a classification scheme for related fauna and flora, as well as models with which to describe populations, or even methods to recognize blood cells or handwriting.

CA is neither a single technique nor a statistical technique. It is a methodology for dividing data into classes, without a preconceived notion of what those classes are, based on relationships within the data. There are many different ways to do this, and some of them use statistical probabilities or statistical quantities such as sum of squares at various points. But overall, the techniques themselves are not really statistical, as they give one no means of assessing the likelihood.

8.9.2. Basic principles

CA methods are based on using heuristic algorithms that build classifications and hierarchies with properties dependent on the implementation of the algorithm. The aim of a CA is the concentration of objects into groups given by a similarity between them. Between the different groups, there should be less similarity than within the groups. There are two main steps in a CA:

- Selection of a measure of proximity (making of distance matrix): the extent of proximity for two objects, each is being measured and expressed by a numerical value,
- Selection of a fusion algorithm: based on their proximities, the objects are grouped into clusters in a way that distinguishes them from other clusters. Further steps result in clusters being clustered iteratively thus building a hierarchy of clusters.

The complete process of generalised hierarchical clustering can be summarised as follows:

1. object selection (X,Y) and definition of a set of variables for each object (x_i,y_i)
2. object standardization
3. computation of the similarities (s) and dissimilarities (d) amongst the objects
4. use of a CA method to create groups of similar entities $D((C_i \cup C_j),C_k)$
5. validation of the resulting cluster solution.

Each of these steps is essential in the use of CA in applied data analysis, and each is briefly outlined below.

8.9.3. Objects selection

Suppose there are n objects to be clustered. The first represents the objects by means of p measurements or attributes. These measurements can be arranged in an n-by-p matrix, equation (8.18), where the rows correspond to the objects and the columns to the measurements.

$$\begin{array}{c} \\ n \text{ objects} \end{array} \begin{array}{c} p \text{ variables} \\ \begin{bmatrix} x_{11} & \cdots & x_{1f} & \cdots & x_{1p} \\ \vdots & & \vdots & & \vdots \\ x_{i1} & \cdots & x_{if} & \cdots & x_{ip} \\ \vdots & & \vdots & & \vdots \\ x_{n1} & \cdots & x_{nf} & \cdots & x_{np} \end{bmatrix} \end{array} \tag{8.18}$$

8.9.4. Interval-scaled variables (Standardizing the data)

Raw data may be recorded at different scales. Inevitably, this can result into data that has very different ranges. These discrepancies can distort the proximity calculations. Standardizing converts the original measurements into unitless variables. There are many ways of achieving this, including converting to normal scores (equation (8.19)), scaling to give a minimum of 0 and a maximum of 1 (equation (8.20)) or scaling from –1 to +1 (equation (8.21)).

The standardized measurements can be defined by:

$$z_{if} = \frac{x_{if} - m_f}{\sigma_f} \tag{8.19}$$

where m_f is the mean value of the variable and σ_f is standard deviation. Sometimes instead of σ_f mean, absolute deviation s_f is used. This measure is more robust in the sense that one outlying observation will not have such a large influence on s_f and is shown as:

$$z_{if} = \frac{x_{if} - \min X_i}{\max X_i - \min X_i} \tag{8.20}$$

$$z_{if} = 2 \cdot \frac{x_{if} - \min X_i}{\max X_i - \min X_i} - 1 \tag{8.21}$$

In equations (8.20) and (8.21) X_i presents a vector of all values in object i.

8.9.5. Similarity and dissimilarity measures

The quantitative estimation of similarity has been dominated by the concept of metrics; this approach to similarity represents cases as points in a coordinate space such that the observed similarities and dissimilarities of the points correspond to metric distances between them. There are four standard criteria that can be used to judge whether a similarity measure is a true metric. These are:

(D1) $d(x,y) \geq 0$	distances are nonnegative numbers
(D2) $d(x,y) = 0$ if $x = y$	distance to itself is zero
(D3) $d(x,y) = d(y,x)$	symmetry of the distance function
(D4) $d(x,z) + d(x,z) \geq d(x,y)$	triangle inequality

Many, but not all, distance measures discussed below are metrics. A number of correlation measures are not metric. There are four types of similarity measures: (1) correlation coefficients, (2) distance measures, (3) association coefficients, and (4) probabilistic similarity coefficients. Each of these methods has advantages and disadvantages that must be considered before a decision is made as to which one to use.

Basically, dissimilarities are nonnegative numbers $d(i,j)$ that are small (close to zero) when i and j are "near" to each other and which become large when i and j are very different. Dissimilarities satisfy (D1), (D2) and (D3), but the main difference with the distances is that (D4) can no longer be relied upon.

Instead of using a dissimilarity coefficient $d(i,j)$ to indicate how remote two objects i and j are, it is also possible to work with a similarity coefficient $s(i,j)$. The more objects i and j are alike (or close), the larger $s(i,j)$ becomes.

8.9.5.1. Distance measures

Because of their intuitive appeal, distance measures have enjoyed widespread popularity. Technically, they are best described as dissimilarity measures; most of the popular coefficients demonstrate similarity by high values within their ranges, but distance measures are scaled in reverse.

The most popular distance measures are:

Euclidean distance $$d(X,Y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Manhattan distance (City Block) $$d(X,Y) = \sum_{i=1}^{n}|x_i - y_i|$$

Minkowski distance $$d(X,Y) = \sqrt[p]{\sum_{i=1}^{n} |x_i - y_i|^p}$$

Mahalanobis $$d(X,Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^T \mathbf{K}^{-1} (x_i - y_i)}$$

Standardized Euclidian distance $$d(X,Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i) \mathbf{V}^{-1} (x_i - y_i)^T}$$

Notice when p=1 in Minkowski distance, it becomes the Manhattan case, and when p=2, it becomes the Euclidian case. Mahalanobis and Standardized Euclidian distances consider the links between the variables and **V** and **K** to be diagonal matrices. Elements of the matrix **V** are the variance of the variables and **K** is the sample covariance matrix.

8.9.5.2. Association coefficients

This type of measure is used to establish similarity between cases described by binary variables.

8.9.5.3. Probabilistic similarity coefficients

They can be used only with binary data.

8.9.6. CA techniques

The choice of a clustering algorithm depends both on the type of available data and on the particular purpose. Sometimes several algorithms are applicable, and a priori arguments may not suffice to narrow down the choice to a single method. Therefore it is good practice to run several algorithms on the same data; this is so because CA is mostly used as a descriptive or exploratory tool, in contrast to statistical tests, which are carried out for inferential or confirmatory purposes.

CA techniques can be roughly "classified" into the following types:

Hierarchical techniques - in which the classes themselves are classified into groups, the process being repeated at different levels to form a tree.

Optimisation-partitioning techniques - in which the clusters are formed by optimisation of a "cluster criterion". The classes are mutually exclusive, thus forming a partition of the set of entities.

Density or mode-seeking techniques – in which clusters are formed by searching for regions containing a relatively dense concentration of entities.

Clumping techniques – in which the classes or clumps can overlap.

Others – methods which do not fall clearly into any of the four previous groups.

Each of these families adds a different perspective to the creation of groups, and the results obtained when different methods are applied to the same data can be very different. It is important to remember that when faced with the difficult choice of which clustering method to use, the method must be compatible with the desired nature of the classification, the variables to be used, and the similarity measure used to estimate the resemblance between cases.

8.9.7. Hierarchical methods

Generally, the hierarchical methods are more widely used because they give greater insight into the overall structure. They generate a hierarchy of partitions by means of a successive merging (agglomerative techniques) or splitting (divisive techniques) of clusters. A divisive method begins with all cases in one cluster. This cluster is gradually broken down into smaller clusters. Agglomerative techniques usually start with single member clusters. These are gradually fused until one large cluster is formed.

Figure 8.12 shows what happens with a data set with $n = 5$ objects. Agglomerative methods (indicated by the upper arrow, pointing to the right) start when all objects are apart (i.e., at step 0 we have n clusters). Then at each step two clusters are merged, until only one is left. On the other hand, divisive methods start when all objects are together (i.e., at step 0 there is one cluster) and at each of the following steps, a cluster is split up until there are n of them. In this example, the agglomerative and divisive hierarchies coincide, but very often they are different.

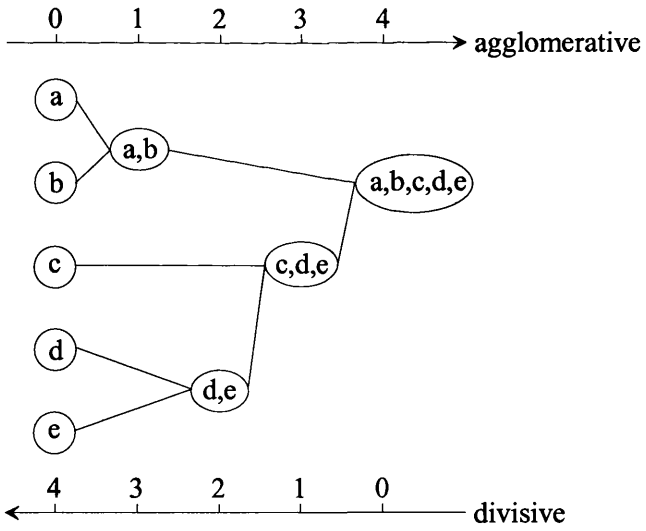


Figure 8.12: Distinction between agglomerative and divisive techniques

Agglomerative methods are based on the algorithm given in Figure 8.13. At the beginning, all objects are apart (n groups). Firstly, we look for groups that have minimal distance $\min(D(C_p,C_q))$ and merge them to a new group $C_s = C_p \cup C_q$. We calculate the measure of dissimilarity D and merge the groups with minimal distance again. The procedure is repeated until all groups are merged to a single group after n-1 steps.

Single linkage clustering (minimum or nearest neighbour method) joins the two clusters that have the two most similar individual points. This can cause many problems; often what one gets is a series of clusters and each has one new member. This is also called chaining.

Complete linkage clustering (maximum or furthest neighbour method) joins the two clusters that have the greatest dissimilarity between two points. This method tends to create "globular-shaped" clusters that have unequal variances and sample sizes. It is a very practical and useful method in terms of being able to straightforwardly interpret the output. This method tends to produce very tight clusters of similar cases.

Average linkage clustering calculates similarities between clusters using average values. Other variants of average linkage are designed to calculate the similarity between the centroids of two clusters that might be merged (*centroid clustering*). The centroid of a cluster is the average point in the multidimensional space defined by the dimension.

Ward's Method uses an analysis of variance approach to evaluate the distance between clusters. In short, this method attempts to minimize the sum of squares of any two (hypothetical) clusters that can be formed at each step.

Rules as to how distances are measured between the clusters are presented in Table 8.1.

Table 8.1: Clustering algorithms

Algorithm	Equation
Single linkage (nearest neighbour)	$D(C_i \cup C_j, C_k) = \min\left(D(C_i,C_k), D(C_j,C_k)\right)$
Complete linkage (furthest neighbour)	$D(C_i \cup C_j, C_k) = \max\left(D(C_i,C_k), D(C_j,C_k)\right)$
Average linkage	$D(C_i \cup C_j, C_k) = \frac{1}{(n_i+n_j)\cdot n_k}\sum_{i=1}^{n_i}\sum_{j=1}^{n_j} D(U,V)$ $U \in C_i \cup C_j \;\; ; \;\; V \in C_k$ n_i, n_j and n_k denote number of objects in clusters
Centroid	$D(C_i \cup C_j, C_k) = D^2(T_{ij}, T_k)$ T_{ij} denotes arithmetical mean of merged cluster $C_i \cup C_j$, while T_k presents arithmetical mean of C_k
Ward	$D(C_i \cup C_j, C_k) = \frac{(n_i+n_j)\cdot n_k}{n_i+n_j+n_k} D^2(T_{ij}, T_k)$

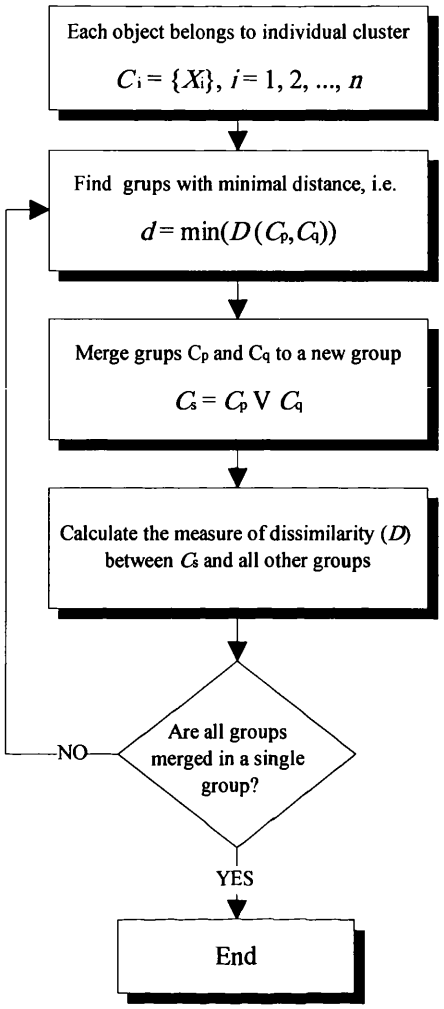


Figure 8.13: Agglomerative method algorithm

Lance and Williams [99] have developed a formula that can be used to describe linkage rules in a general form for any hierarchical agglomerative method. This rule is defined as:

$$D(C_i \cup C_j, C_k) = \alpha_1 D(C_j, C_k) + \alpha_2 D(C_i, C_k) + \beta D(C_i, C_j) + \gamma \left| D(C_i, C_k) - D(C_j, C_k) \right| \quad (8.22)$$

$C_i \cup C_j$: new cluster, C_i, C_j: merged clusters,

C_k: remaining cluster, α_1, α_2, β and γ: parameters

The parameters (α_1, α_2, β and γ) and properties for the methods presented above are described in Table 8.2.

In some cases, the defined grouping boundaries encounter some modifications: if remote elements are being added to a cluster, the cluster space is contracted (contraction effect), as shown in Figure 8.14. Chaining is where single samples join a larger cluster each time, so that instead of a CA, one really has an ordination, i.e., there is no true hierarchical structure. A large

amount of small and very small groups result in a stretching of the cluster space (dilatation effect), as shown in Figure 8.15. Conservative methods try to avoid such effects. Reversals are caused when an entity joins another cluster at a lower level of similarity than was there before, making the hierarchical structure go "backwards" in some cases, as shown in Figure 8.16.

Hierarchical clustering solutions are typically represented by a dendrogram, Figures 8.14-8.16. A dendrogram is a rooted weighted tree, with leaves corresponding to elements (X,Y). Each edge defines the cluster of elements contained in the subtree below that edge. The edge's weight (or length) reflects the dissimilarity between that cluster and the remaining elements.

Table 8.2: Agglomerative methods and their properties

Method	α_1	α_2	β	γ	**Monotone**	**Influence on the cluster space**
Single linkage	$\frac{1}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	yes	contraction; building of chains
Complete linkage	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	yes	dilatation; tendency of small groups
Average linkage	$\frac{n_i}{n_i+n_j}$	$\frac{n_j}{n_i+n_j}$	0	0	yes	conservative
Centroid	$\frac{n_i}{n_i+n_j}$	$\frac{n_j}{n_i+n_j}$	$-\frac{n_i+n_j}{(n_i+n_j)^2}$	0	no	*
Ward	$\frac{n_i+n_k}{n_i+n_j+n_k}$	$\frac{n_j+n_k}{n_i+n_j+n_k}$	$-\frac{n_j+n_k}{n_i+n_j+n_k}$	0	yes	*

* Influence on the space depends on $\alpha 1$, $\alpha 2$, β and γ

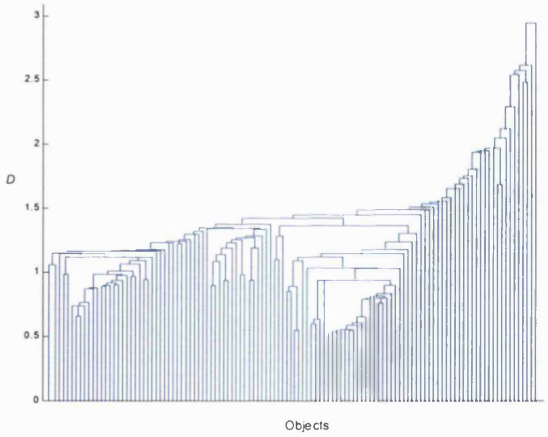


Figure 8.14: Contraction effect

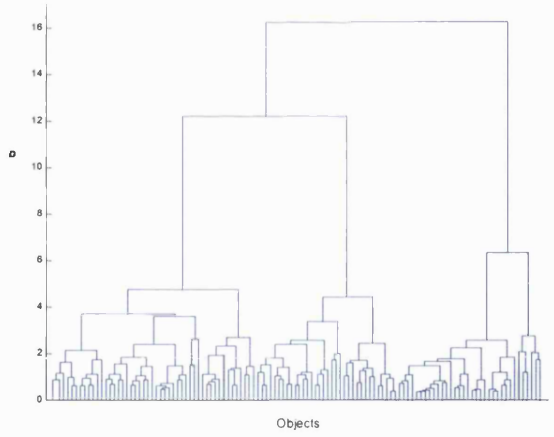


Figure 8.15: Dilatation effect

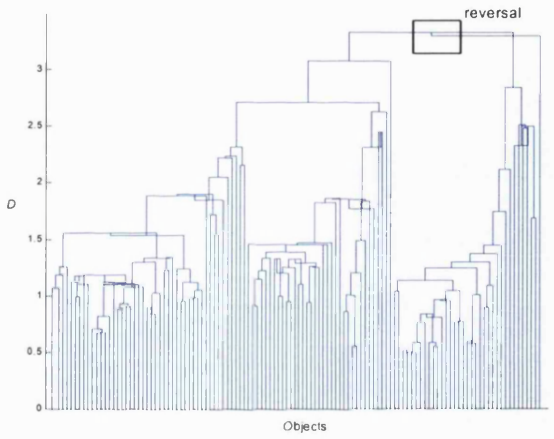


Figure 8.16: Reversal effect

8.9.8. Assessment of solution

Once the objects have been clustered, there is a need to compare the various clusters to get an idea of how they differ. A straightforward, simple approach is to compare the clusters with respect to their means and variances on the p variables used in computing the inter-object similarities or on a set of external variables for which information is available on the cluster members, but which was not used in the clustering procedure.

One way to determine the natural cluster divisions in a dataset is to compare the length of each link in a cluster tree with the lengths of neighbouring links below it in the tree. If a link is approximately the same length as neighbouring links, it indicates that there are similarities between the objects joined at this level of the hierarchy. These links are said to exhibit a high level of consistency. In the hierarchical cluster tree, the dataset may naturally align itself into clusters. This can be particularly evident in a dendrogram diagram where groups of objects are

densely packed in certain areas and not in others. The maximum difference between distances in dendrogram can also be the measure for decision as to how many clusters should be taken.

Validating the quality of a cluster solution can be achieved by statistical testing; for example, the maximum value of the *F*-statistic that is obtained among all alternative groupings [101].

8.9.9. An example of hierarchical clustering

The operation of a hierarchical clustering algorithm is illustrated in Figure 8.17 using two-dimensional data set. The hierarchical clustering yields a dendrogram representing the nested group of samples and similarity levels at which grouping changes. The resulting dendrogams (Figure 8.18) for two linkage algorithms, Ward and Average, respectively, correspond to 29 points as shown in Figure 8.17. A comparison of both dendrograms indicates that mergers take place at different distances, because distance is defined differently for each case (see Table 8.1). The dendrogram can be cut off at different distances to yield different clustering of the data. What is the optimal number of clusters? With reference to section 8.9.8, the optimal cut off of the vertical lines in the dendrogram is where the longest link (distance) is between two successive clusters. Determination of the final cluster number based on the distance between two successive clusters from the dendrogram is practically impossible when the number of samples is high. Therefore, better presentation of choosing the appropriate number of clusters is obtained by a graph showing the dependence of distance between two successive clusters with respect to the number of clusters, as given in Figure 8.19. The optimal number of clusters obtained by the Ward linkage method is three, while for the Average linkage method; the optimal number of clusters is five. Dendrograms in Figure 8.18 contain the horizontal line representing the cut off. The Average linkage method gives the result of three well defined clusters and two separate samples with their own clusters, as shown in Figure 8.20. The Ward linkage method on the other hand, suggests three clusters without any separate samples since they are included into the closest cluster. Therefore one can conclude that the Average linkage algorithm is appropriate for separation of particular samples.

When choosing the number of clusters by cutting the dendrogram, the maximum difference between distances in the dendrogram is not always the best solution, e.g. the procedure suggests a small number of clusters having a large scatter. In such a situation the next highest distance between clusters should be taken.

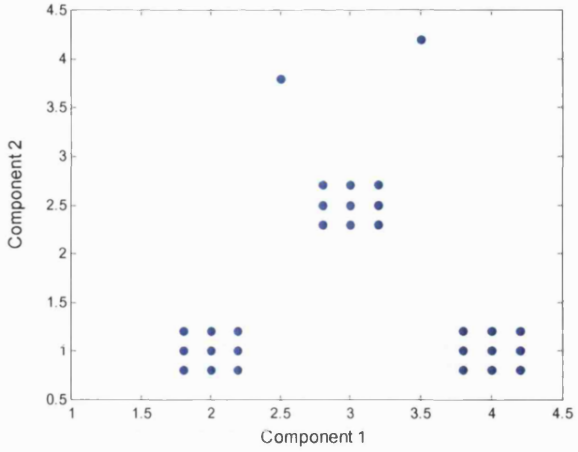


Figure 8.17: The input data where input vectors consists of three components

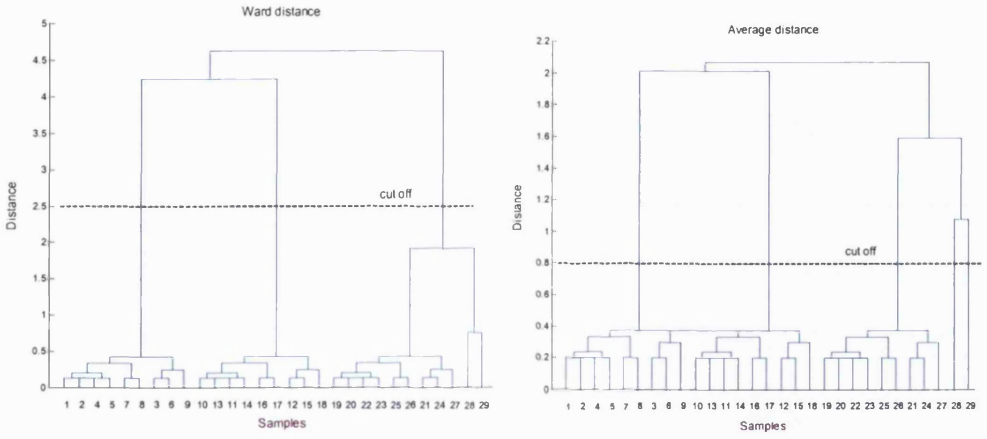


Figure 8.18: Dendrograms based on the Euclidian distance and the Ward (left) and the Average (right) linkage method

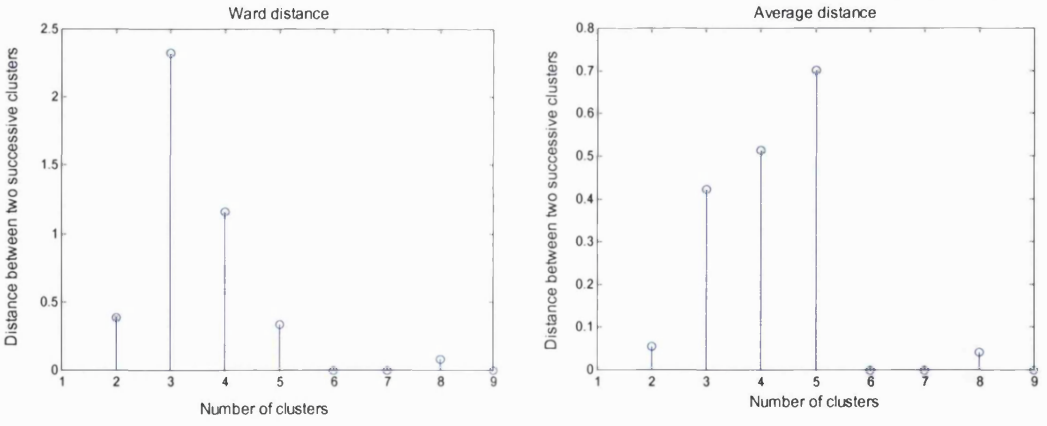


Figure 8.19: The distance between two successive clusters for the Ward (left) and the Average (right) linkage method

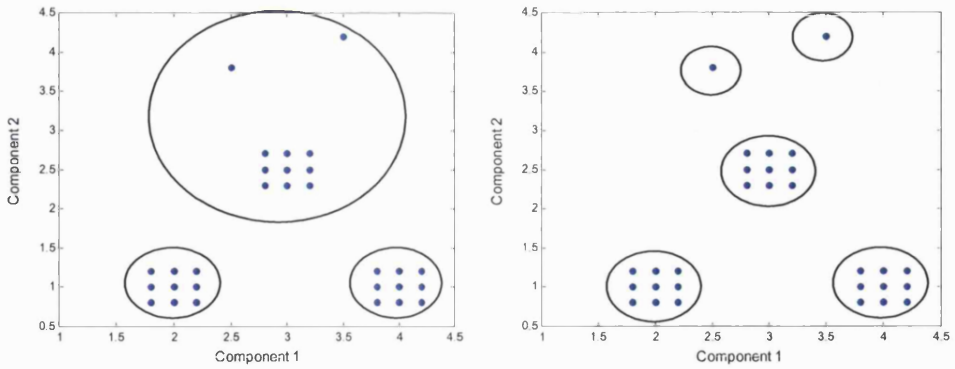


Figure 8.20: Final clusters obtained by Ward (left) and average (right) linkage method

8.10. SUMMARY

A methodology known as DM has been introduced in this chapter. There are many DM techniques, particularly specialized methods for particular types of data and domain. Two methods for dimensionality reduction known as PCA and method of moments has been described. Additionally, this chapter deals with two methods, CA and SOM, that will be used herein for finding groups in captured PD data.

9. DATA MINING ON THE PD DATA USING THE SOM

9.1. INTRODUCTION

The proposed approach based on the SOM was applied for the analysis and interpretation of PD signals. This chapter presents a feasibility study of the proposed approach conducted on signals captured at the Northfleet substation, National Grid Tranco's. These records actually contain a lot of valuable information considering the nature of captured signals (e.g. discharge in the air, discharge in the oil, etc.). If this information can be extracted, comprehended and visualised, the grouping of obtained signals can be enhanced. The data mining approach presented in the previous chapter can be employed to "mine" the hidden information or "knowledge" from the signal and display it in a user-friendly format. The feasibility of the SOM for the grouping of PD signals will be investigated in subsequent chapters of this thesis.

The fundamental concepts of proposed approach are based on a DM method, which has been previously introduced in Section 8.8 and it comprises of various stages, as illustrated in Figure 9.1.

9.2. A FEASIBILITY STUDY BASED ON THE PROPOSED APPROACH

The feasibility study is based on 5200 records captured by antennas and KEMA probe at the Northfleet substation on Saturday 25th of May 2002 (see Chapter 7 for details). Since the signals obtained by KEMA probe serve as a reference for the signals captured by antennas, grouping of the signals using SOM is implemented for signals captured by KEMA probe.

All the signals have to be pre-processed before being submitted for analysis using the SOM, as shown in Figure 9.1. Firstly, each signal is transformed into the frequency domain. The next step is the employment of the scaling method. In this work, the "range" method is applied. It transforms the minimum and maximum values of each component of the input data into "0" and "1" respectively; all other values in each component of the input data are scaled according to equation (8.20). This is to ensure that all normalised features have equal influence in the PCA.

The PCA is then applied. The main purpose of the PCA is to reduce the dimensionality of the data set while retaining the maximum possible variation in the data set. PC's are linear transformations of the original set of variables which are uncorrelated and ordered so that the first few PC's contain most of the variations in the original data set. PCA reduces the dimensionality because most of the dataset are not orthogonal. Equivalently, there exist co-relationships between the column vectors. The number of principal components for further implementation of SOM is chosen with regard to the criterion shown in Section 8.7.3 and the graph shown in Figure 9.2. Sixth eigenvalue has a cumulative variance proportion reaching 80%, (R in equation (8.4)) and therefore the dimension of input data is reduced from (5200 x 5000) to (5200 x 6). The scree plot is shown in Figure 9.3.

Finally the "range" scaling on selected data transformed in PC's is applied again before using the SOM.

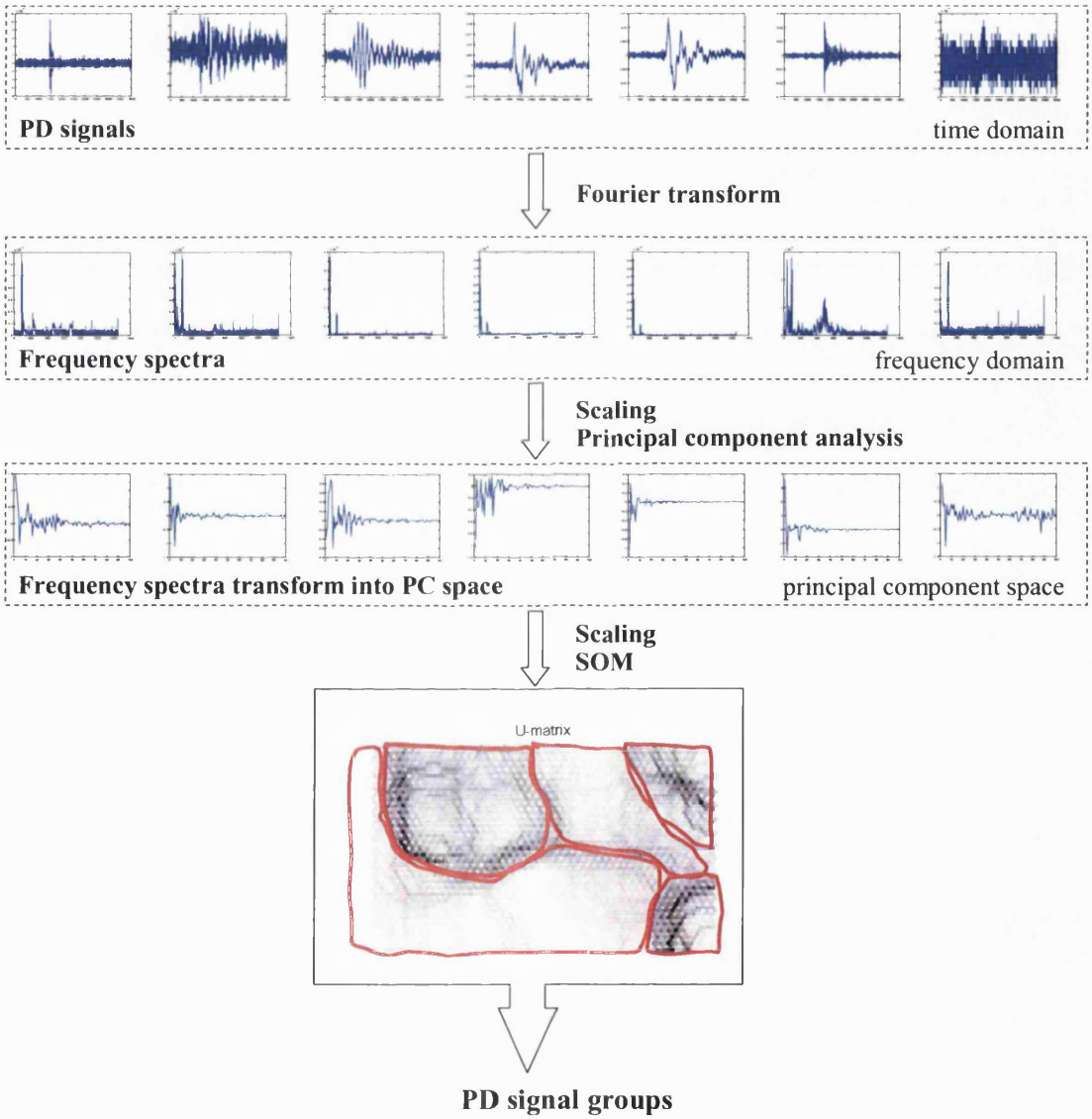


Figure 9.1: Various stages of the proposed approach

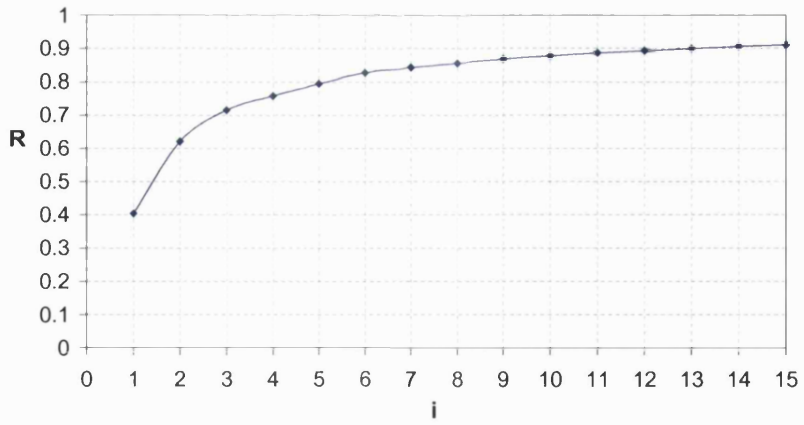


Figure 9.2: The R measure based on *i* principal components

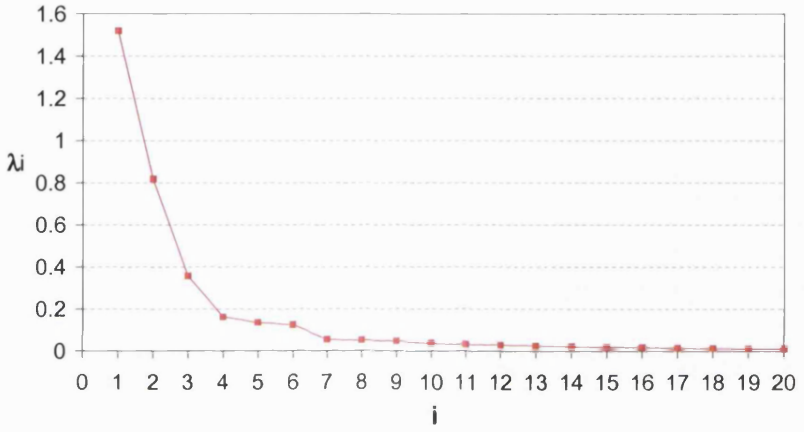


Figure 9.3: The scree plot

In addition, suitable configuration and parameters for the training of SOM is selected according to the guidance for practical application of SOM outlined in Section 8.8.1.4, as summarised in Table 9.1. The map size is (46 x 23) neurons, number of training iterations is 104000, i.e. 20 passes of all training vectors, and SOM quality measure during the fine-tuning phase, such that ε_{AQ} is 0.0655 and ε_T is 0.0506.

The SOM of PD data can be visualised using u-matrix and component-plane illustrations. Component planes give information about the relationships of different data vectors within that specific area. By combining the component planes, a comprehensive view of the whole data set is obtained and this is referred to as the u-matrix.

Optimum SOM of the training data is illustrated in Figure 9.4. By displaying the u-matrix as a grey-level image, the relative distances between adjacent map units on the whole map can be seen. The lighter areas on the map correspond to clusters while the dark lines are borders between clusters.

Colour scale is utilised for the illustration of component planes. The colour bar on the right-hand-side of every component plane illustrates the magnitude of scaled values that are learned by the SOM. The lighter the colour is, the smaller is the relative component magnitude.

Table 9.1: Configuration, training and visualisation parameters of SOM

No.		Phase 1	Phase 2
		Rough Ordering	Fine Tuning
	Configuration of SOM		
1	Shape	Rectangular	
2	Arrangement of neurons	Hexagonal configuration	
3	Number of neurons	$20 \times \sqrt{dlen}$	
4	[x-*dim*, y-*dim*]	$\frac{x-dim}{y-dim} = \sqrt{\frac{\alpha}{\beta}}$	
	Training Parameters		
5	Neighbourhood function	Rectangular	
6	Training method	Sequential	
7	Radius	Max([x-*dim*, y-*dim*])	1/8 × Radius_Phase 1
8	Learning rate	0.05	0.001
9	Learning rate function	Linear	
10	Training iterations	According to Section 8.8.1.4	

dlen: Total number of input vectors.
x-dim, *y-dim*: Number of neurons in the x- and y-dimension of SOM.
α, β: First and second eigenvalues of the covariance matrix of input data.

9.3. ANALYSIS ON REVEALED FEATURES OF THE PD DATA

From Figure 9.4, it is evident that several clusters are legibly formed in the u-matrix. Since the exact number of clusters in the described case is not known, eleven clusters are considered. The number eleven is chosen empirically.

Using the component plane visualisation, the contents of each cluster given in Figure 9.5 can be interpreted as if they correspond to different types of PD signals presented to the SOM. For example, in group '1', Components 1 and 2 are extremely large, Component 3 is very small, while Components 4, 5 and 6 are medium (see component planes in Figure 9.4).

Figures 9.6 to 9.16 contain six graphs. Graph a) presents PC values of signals belonging to a certain group. Each signal (measurement) is presented by six numbers belonging to the first six PC's. Abscissa contains all signals (measurements) belonging to the group and the ordinate contains the value of PC. Graph b) presents the mean value of frequency spectrum for windows 0-100MHz, 100-200 MHz, 200-300 MHz, 300-400 MHz, 400-500 MHz, 500-600 MHz and 600-700 MHz, respectively. The average values of the seven frequency spectra windows are

chosen to give the information on frequency content of the PD signals. Note, that for PC values, the actual, non-scaled, values are given. Graph c) presents a typical PD signal obtained by the KEMA probe and its frequency spectrum (graph d). Figures 9.6 to 9.16 also include PD signals captured by antenna (graphs e), and its frequency spectrum (graph f). When comparing graphs c) to f), it must be emphasised that ordinate axis of signals of different groups as well as frequency spectra are not on the same scale.

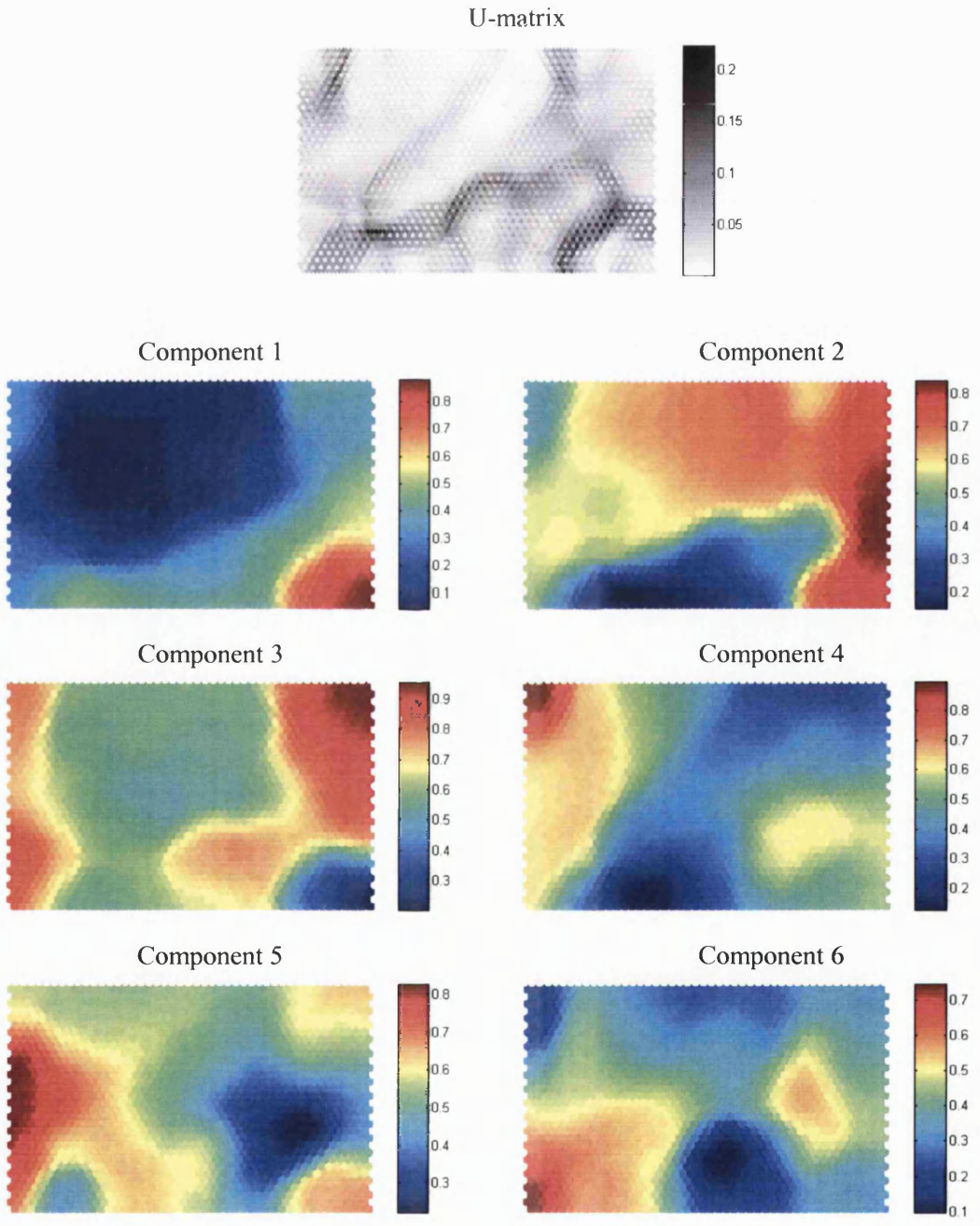


Figure 9.4: U-matrix and component planes of the SOM

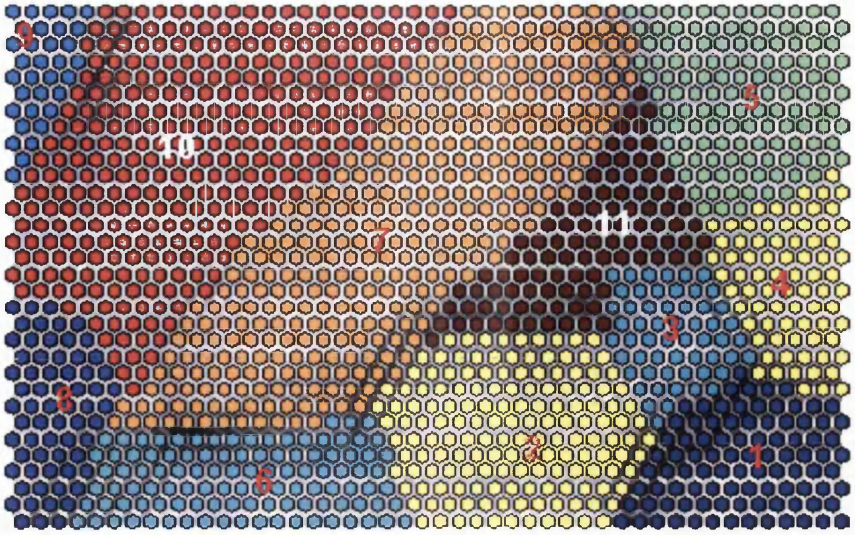
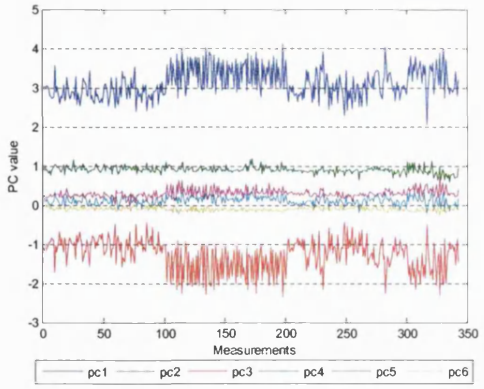
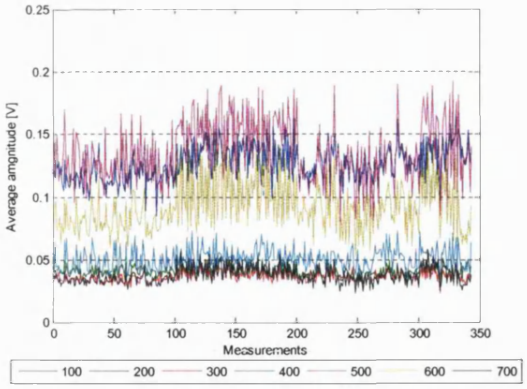


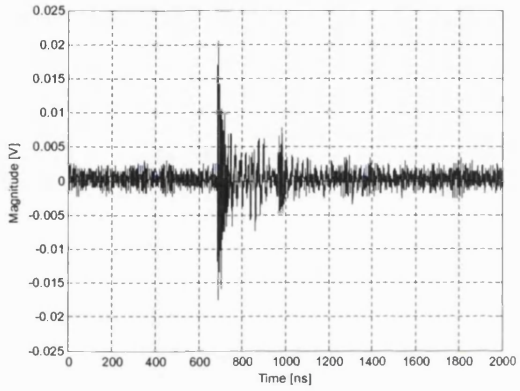
Figure 9.5: Chosen clusters assigned in the u-matrix



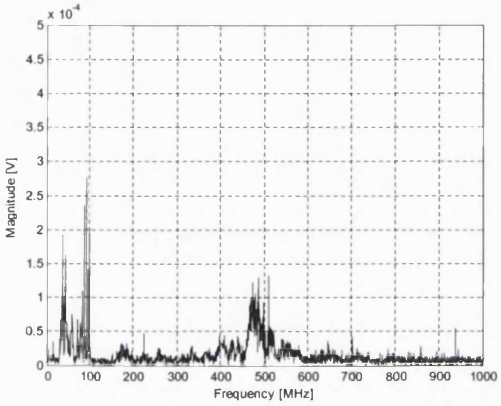
a) PC values of PD signals



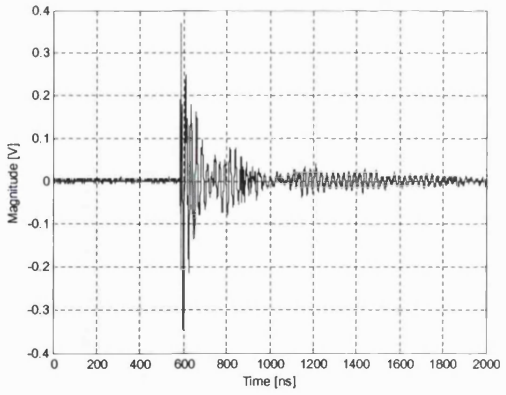
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



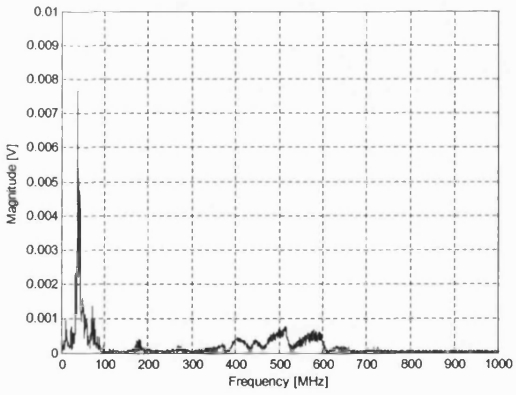
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe

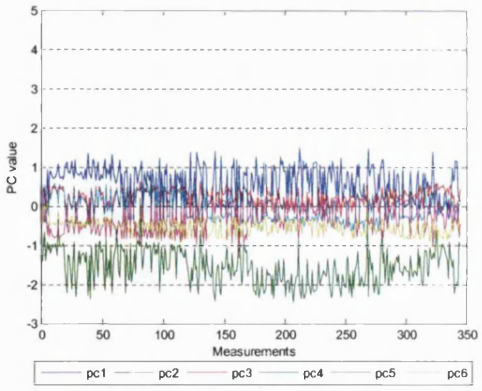


e) Typical signal in time domain obtained by antenna

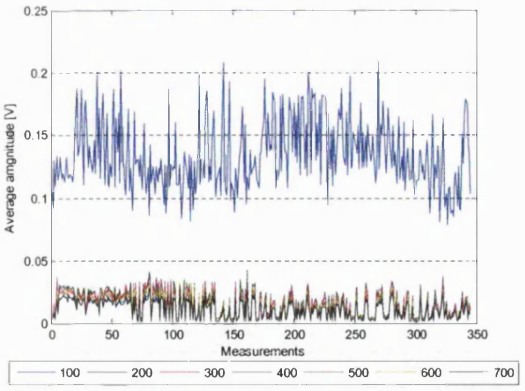


f) Frequency spectrum of the signal obtained by antenna

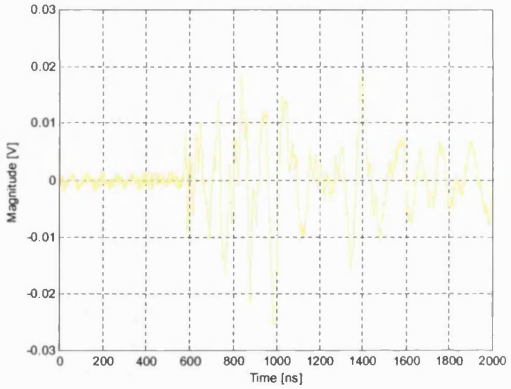
Figure 9.6: Group 1



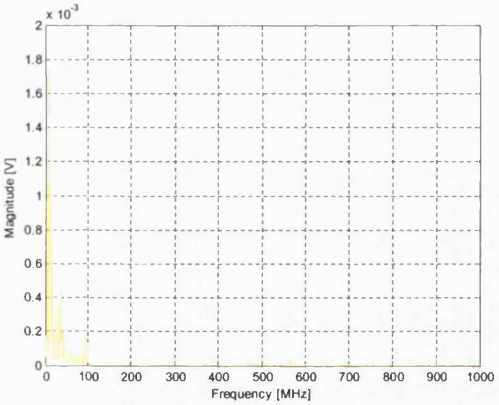
a) PC values of PD signals



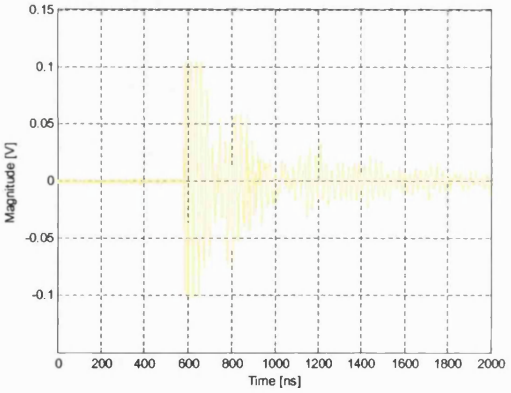
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



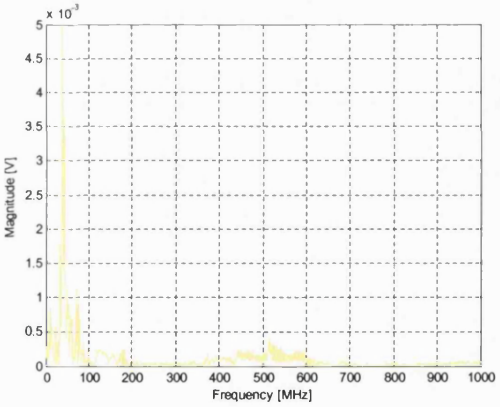
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe

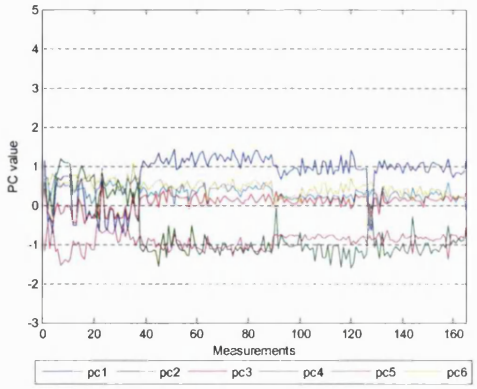


e) Typical signal in time domain obtained by antenna

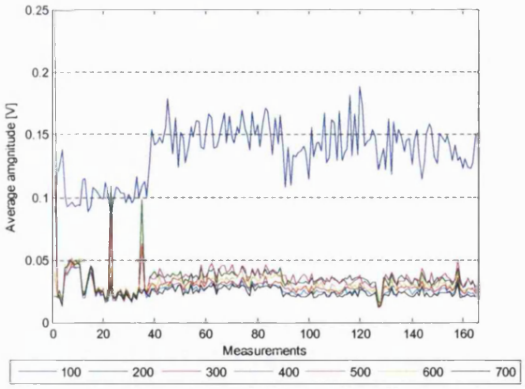


f) Frequency spectrum of the signal obtained by antenna

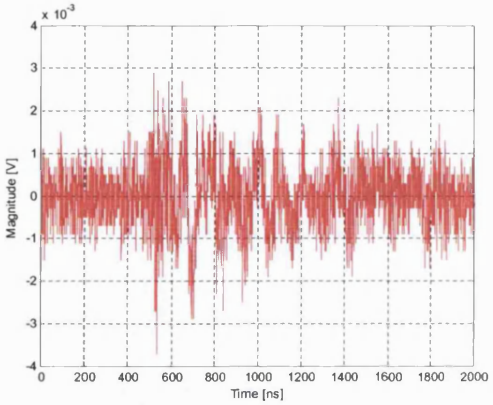
Figure 9.7: Group 2



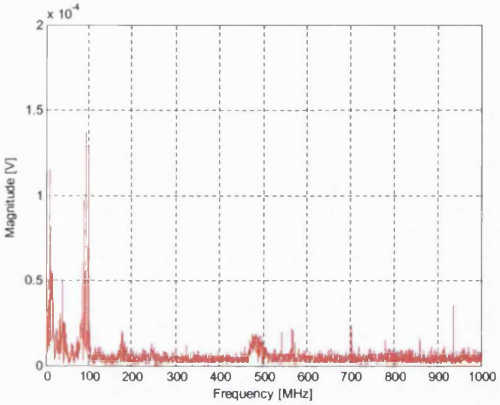
a) PC values of PD signals



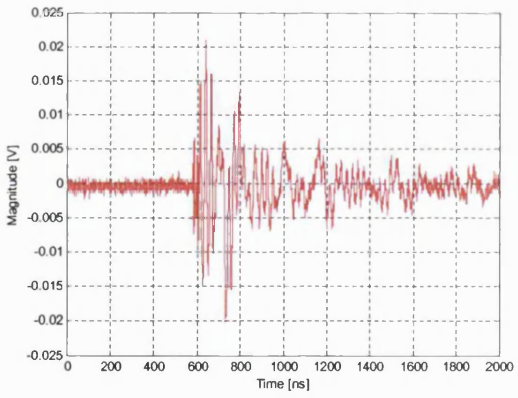
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



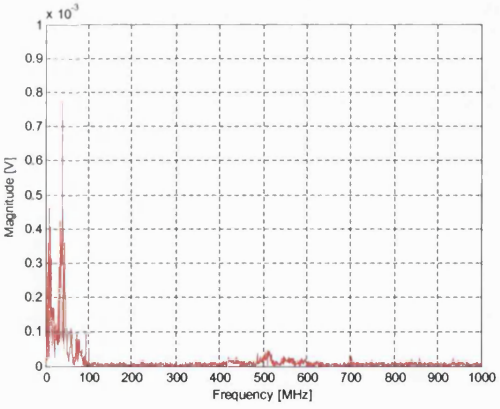
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe

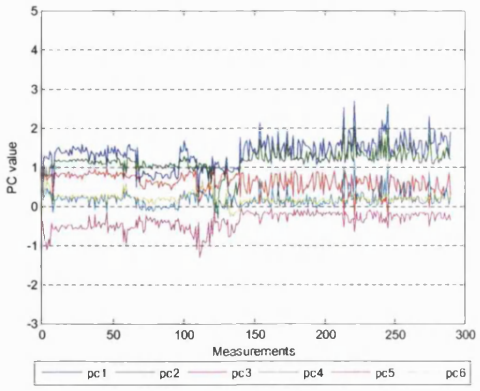


e) Typical signal in time domain obtained by antenna

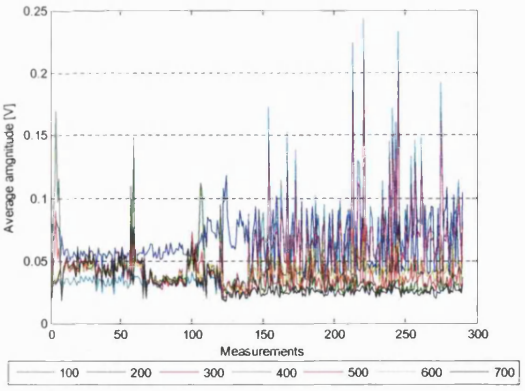


f) Frequency spectrum of the signal obtained by antenna

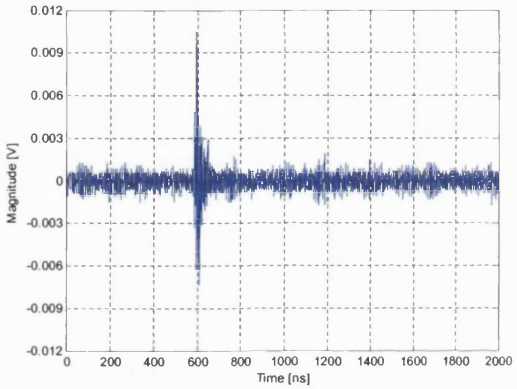
Figure 9.8: Group 3



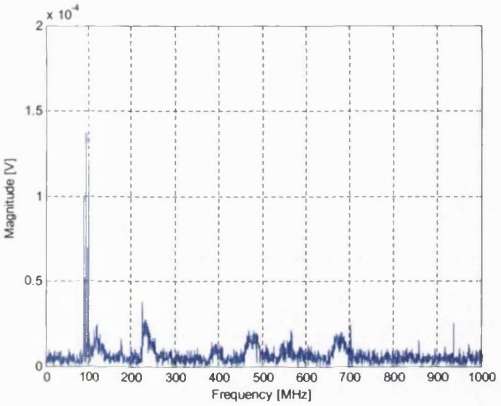
a) PC values of PD signals



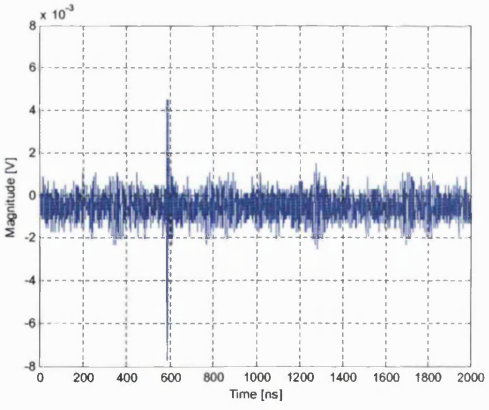
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



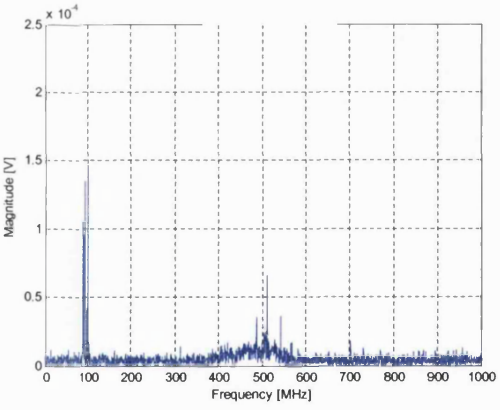
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe

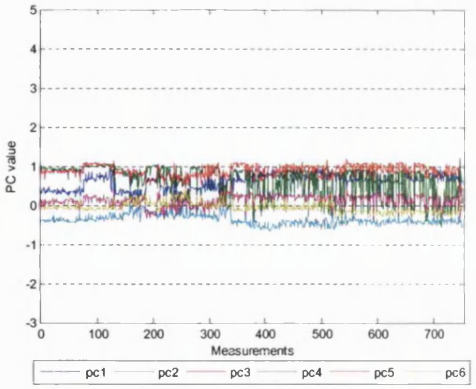


e) Typical signal in time domain obtained by antenna

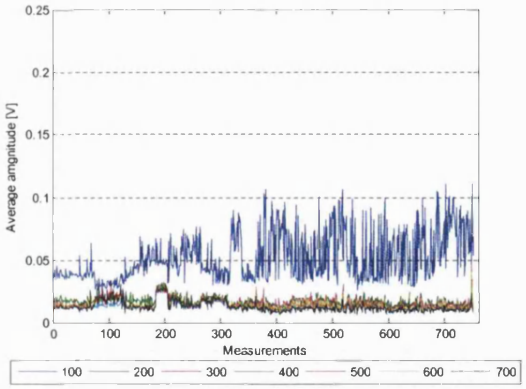


f) Frequency spectrum of the signal obtained by antenna

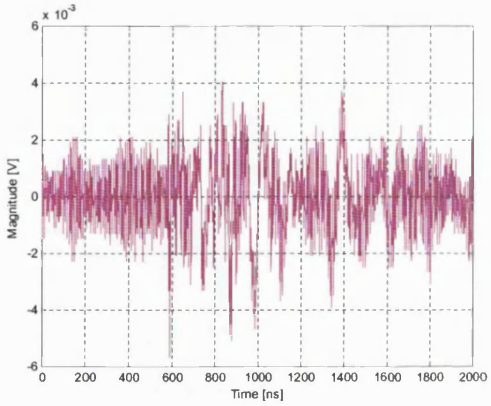
Figure 9.9: Group 4



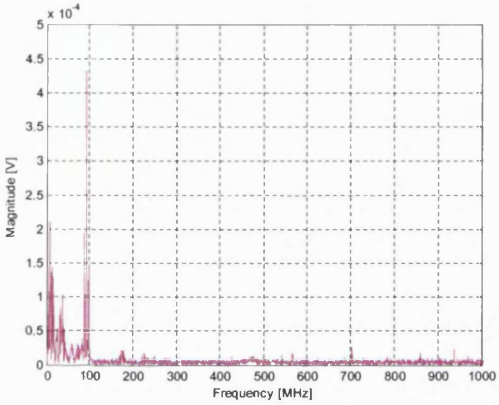
a) PC values of PD signals



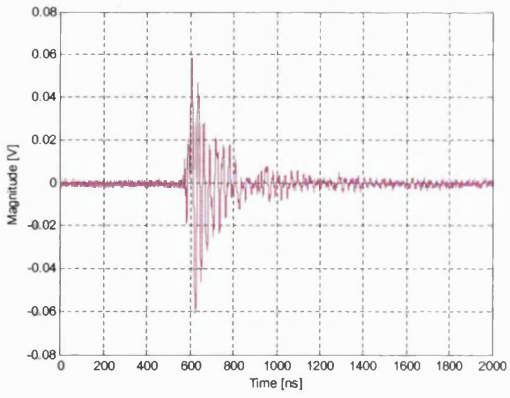
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



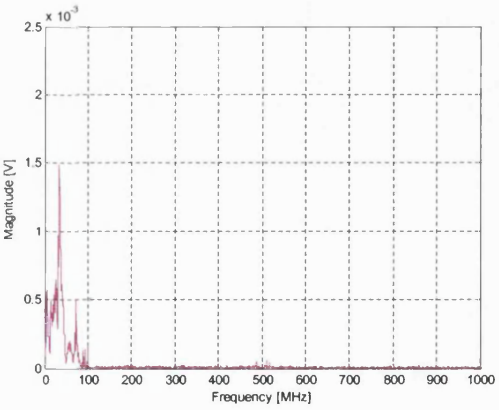
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe

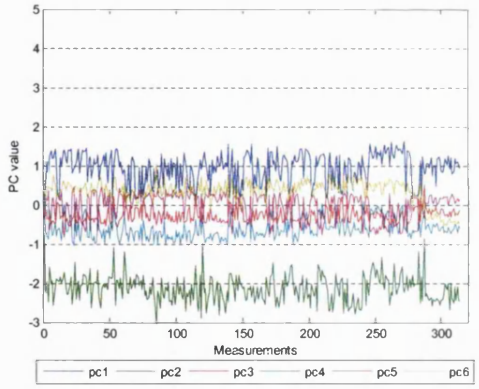


e) Typical signal in time domain obtained by antenna

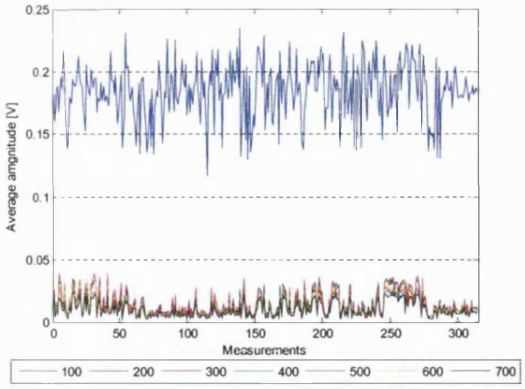


f) Frequency spectrum of the signal obtained by antenna

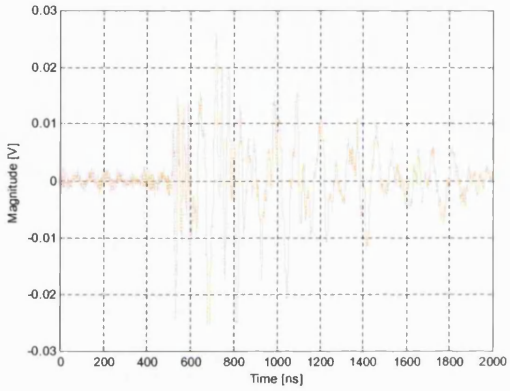
Figure 9.10: Group 5



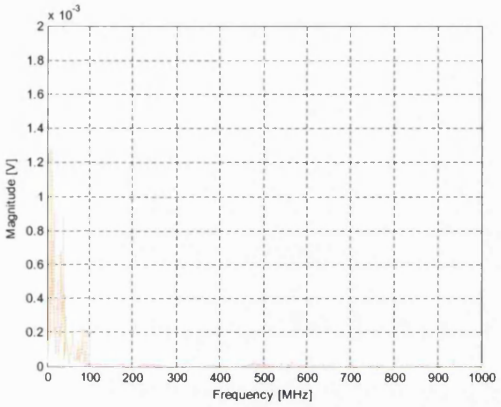
a) PC values of PD signals



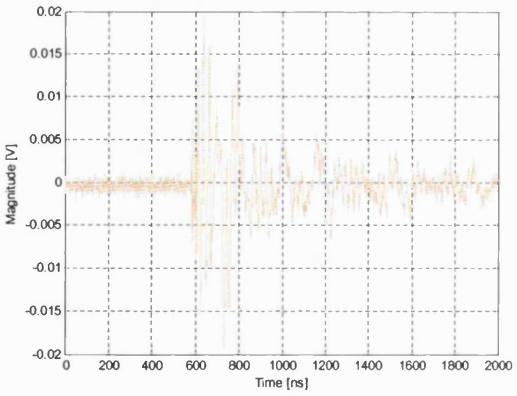
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



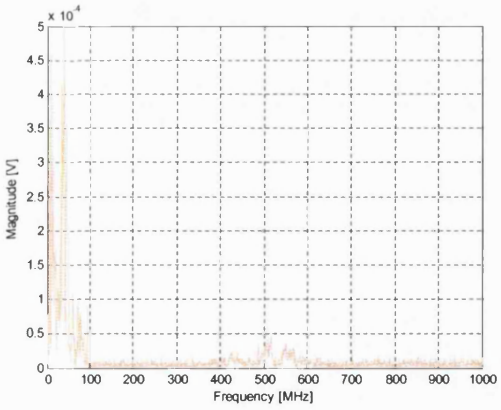
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe

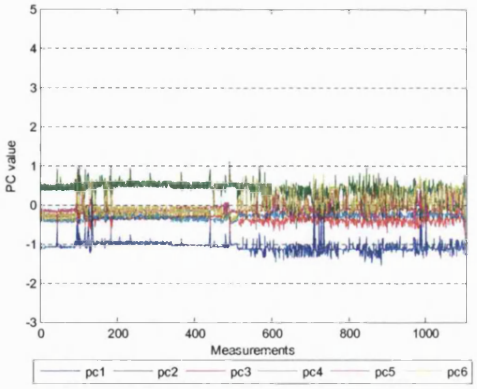


e) Typical signal in time domain obtained by antenna

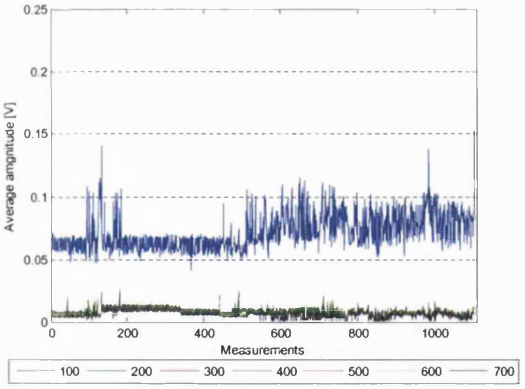


f) Frequency spectrum of the signal obtained by antenna

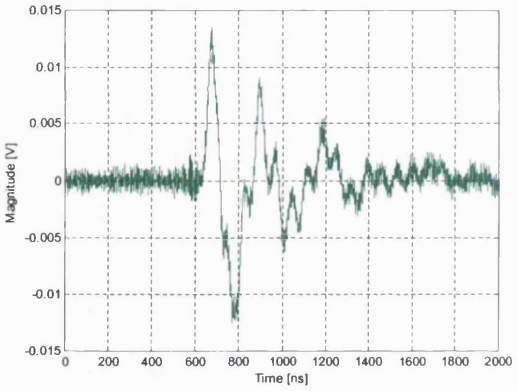
Figure 9.11: Group 6



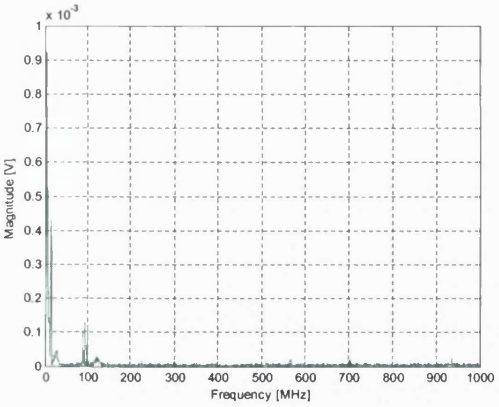
a) PC values of PD signals



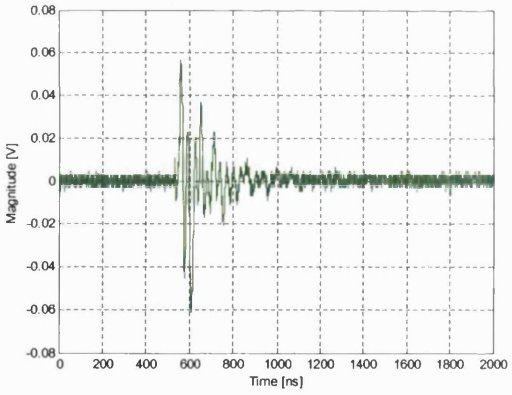
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



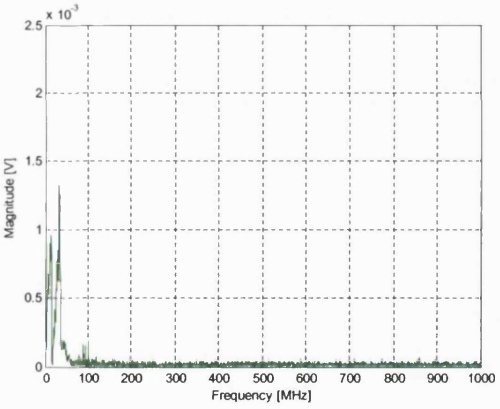
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe

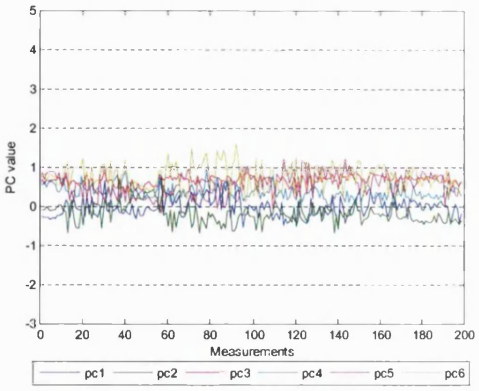


e) Typical signal in time domain obtained by antenna

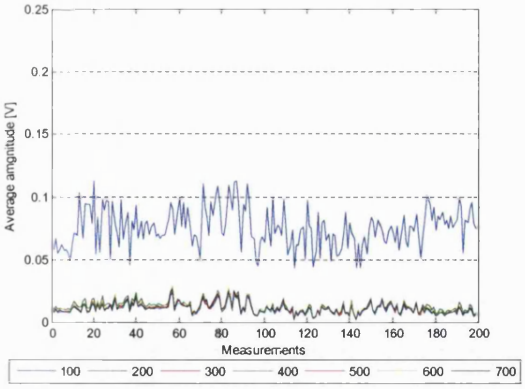


f) Frequency spectrum of the signal obtained by antenna

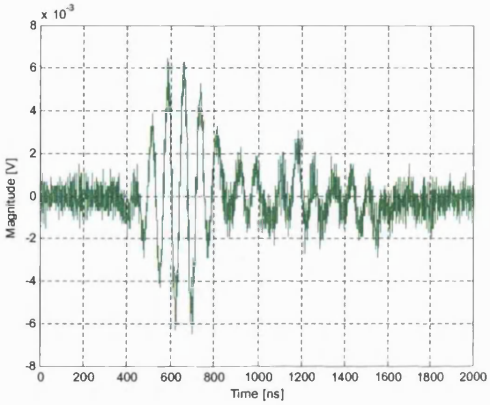
Figure 9.12: Group 7



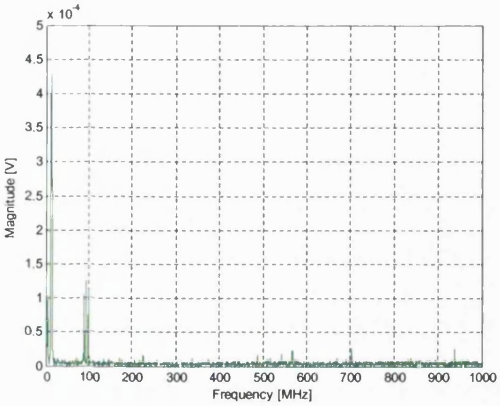
a) PC values of PD signals



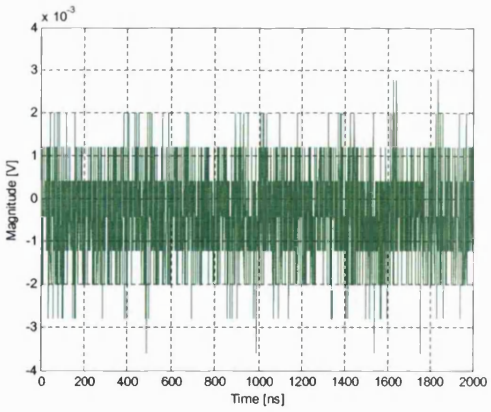
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



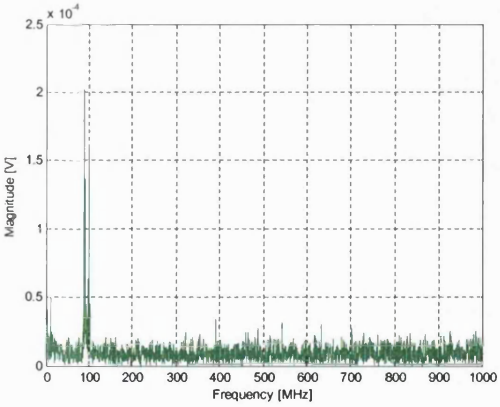
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe

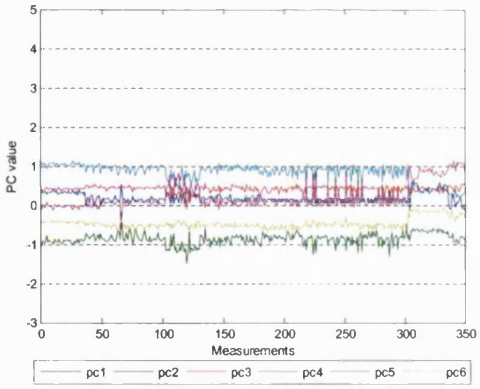


e) Typical signal in time domain obtained by antenna

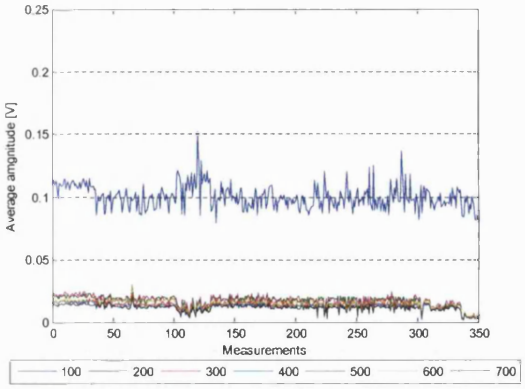


f) Frequency spectrum of the signal obtained by antenna

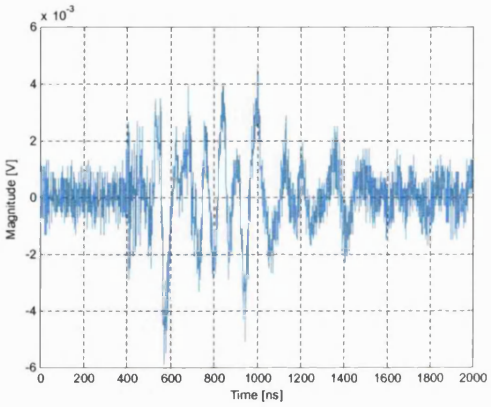
Figure 9.13: Group 8



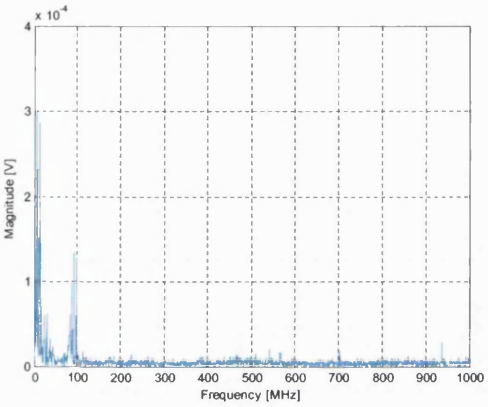
a) PC values of PD signals



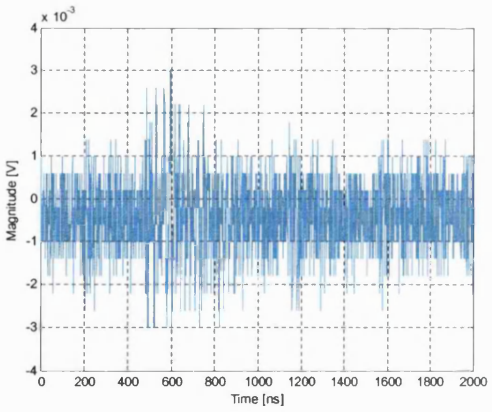
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



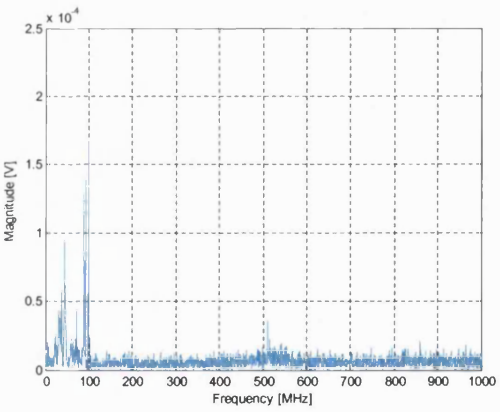
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained the by KEMA probe

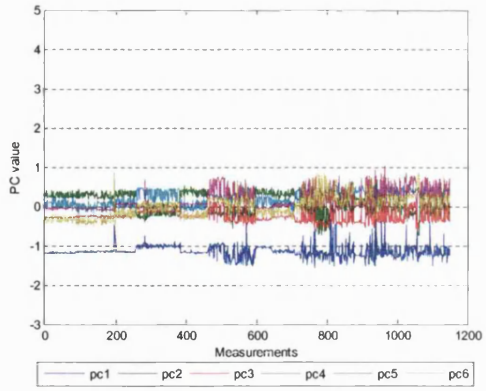


e) Typical signal in time domain obtained by antenna

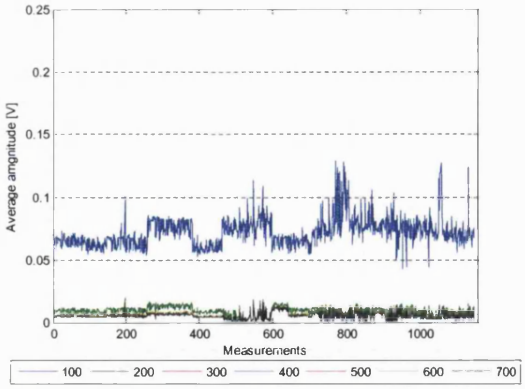


f) Frequency spectrum of the signal obtained by antenna

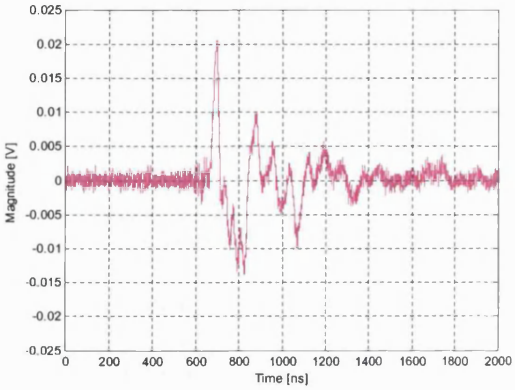
Figure 9.14: Group 9



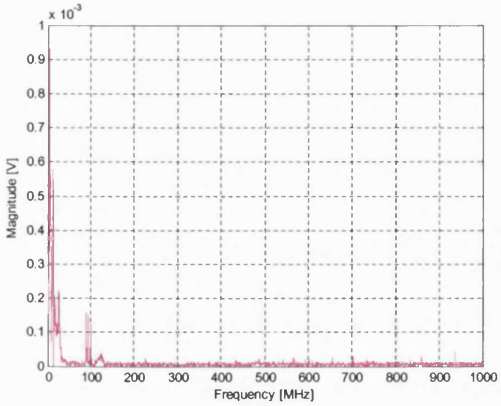
a) PC values of PD signals



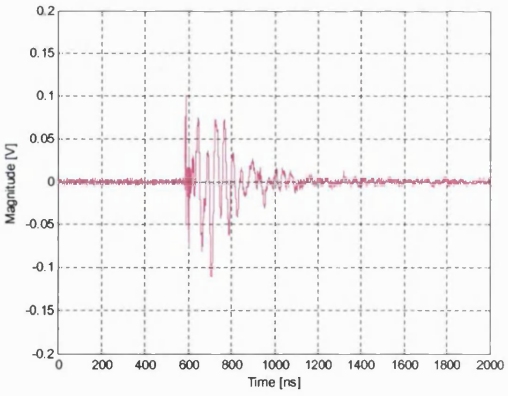
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



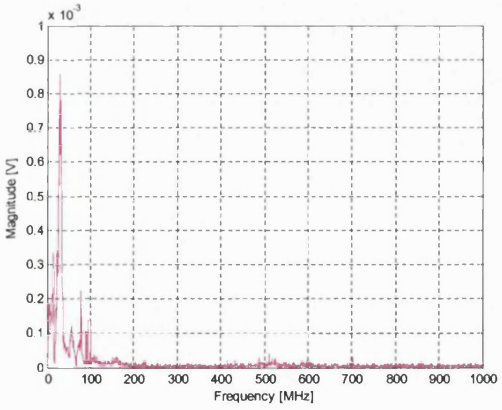
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe

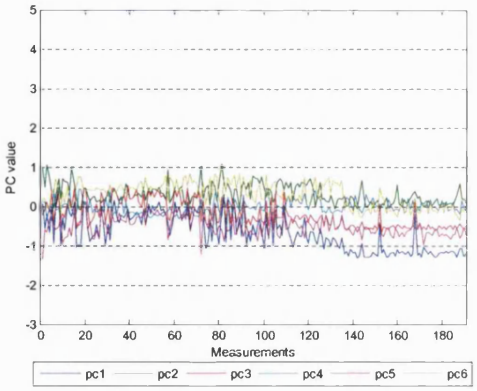


e) Typical signal in time domain obtained by antenna

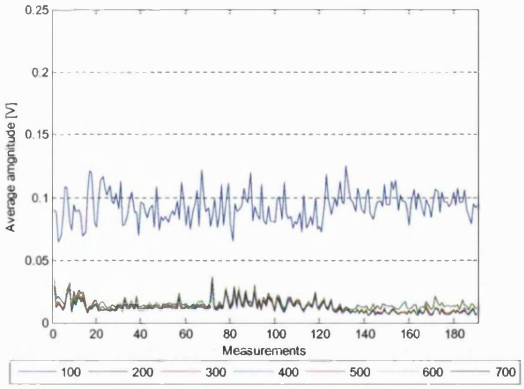


f) Frequency spectrum of the signal obtained by antenna

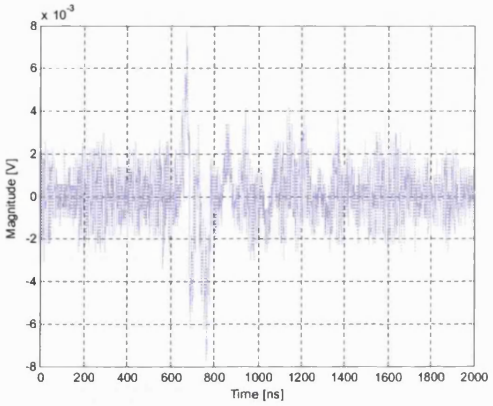
Figure 9.15: Group 10



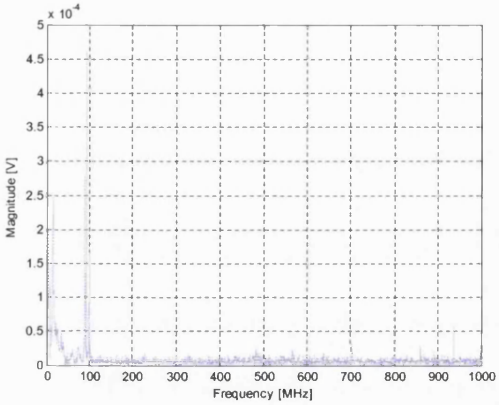
a) PC values of PD signals



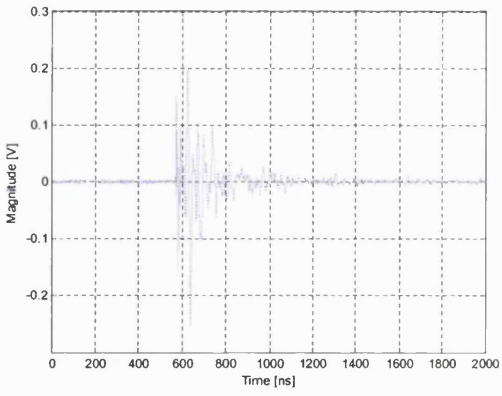
b) average magnitude of frequency spectra of PD signals for 100 MHz windows



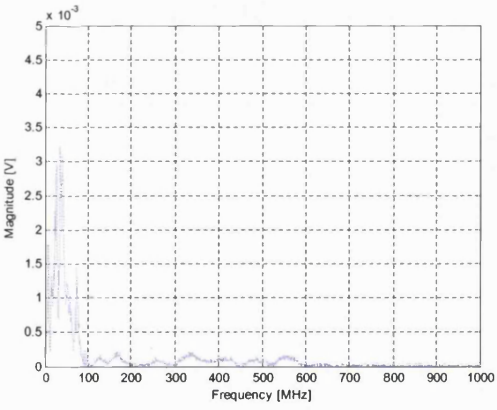
c) Typical signal in time domain obtained by the KEMA probe



d) Frequency spectrum of the signal obtained by the KEMA probe



e) Typical signal in time domain obtained by antenna



f) Frequency spectrum of the signal obtained by antenna

Figure 9.16: Group 11

The following conclusions can be drown from the foregoing observations:

- Group 1 contains 343 captured signals. It can be seen from Figure 9.6.a, that the PC 1 has the highest magnitude and PC 3 the lowest compared to the other groups. The other four components attribute medium values. With regard to the average magnitude frequencies (Figure 9.6.b), frequencies between 400 and 500 MHz are of the highest magnitude. Additionally, all other frequencies have an average magnitude higher than 0.04 V. The frequency spectrum of the signal captured by the KEMA probe (Figure 9.6.d) contains high value of frequencies between 400 and 600 MHz; this is similar to the signal captured by the antenna (Figure 9.6.f).
- Group 4 contains 290 captured signals. It can be seen from Figure 9.9.a, that the PC's 1 and 2 have values higher than 1, while component 3 takes a value between 0 an 1. The other components take medium values. With regard to the average magnitude frequencies (Figure 9.9.b), values are lower compared to Group 1, but relatively high compare to the other nine groups. Signals captured by the KEMA probe and antenna also contain relatively high frequencies above 100 MHz as depicted in Figures Figure 9.9.d and Figure 9.9.f.
- The other signals shown for groups 2-3 and 5- 11 take PC values between -2 and 1, as shown in Figures 9.7-9.8 and 9.10-9.16. In all cases, the dominant average frequency is the window between 0 and 100 MHz, while other windows take magnitudes below 0.04 V. Signals captured by both the KEMA probe and antenna contain dominant frequencies below 100 MHz.
- Further inferences drawn from these signals are: groups 1 and 4 (Figures 9.6 and 9.9) contain PD signals related to discharges in oil by virtue of the fact that the dominant frequency components in the frequency spectra of these signals are above 200 MHz. Furthermore, group 1 consists of PD signals of higher magnitude, i.e. 0.02 V, while group 4 consists of PD signals of magnitude smaller than 0.01 V.
- Groups 2, 3, and 5 to 11 (Figures 9.7 -9.8 and 9.10- 9.16) contain PD signals related to different discharges in air, due to the fact that the dominant spectral content of the signals is below 100MHz.

10. DATA MINING ON THE PD DATA USING THE CA

10.1. INTRODUCTION

A key step in the analysis of PD data is the detection of groups of PDs that exhibit similar expression patterns. This translates to the algorithmic problem of clustering. The goal is to partition the elements into subsets, which are called clusters, so that two criteria are satisfied: homogeneity - elements in the same cluster are highly similar to each other; and separation - elements from different clusters have low similarity to each other.

This chapter deals with the application of clustering to PD data. CA is focused on the hierarchical method. Two distance measures, Euclidian and Mahalanobis, described in Section 8.9.5 and two clustering algorithms, Average and Ward, described in Section 8.9.7 are implemented for PD data mining. Results are represented by a dendogram. Moreover, a comparison of the results obtained by different algorithms is made.

The study is based again on 5200 records captured by antennas and the KEMA probe at the Northfleet substation on Saturday 25th of May 2002. As in the previous DM method, signals obtained by the KEMA probe serve as a reference for signals obtained by antennas. Therefore, grouping of signals using CA is implemented for signals captured by the KEMA probe.

Again, all signals have to be pre-processed before being submitted for analysis using the CA, as shown in Figure 10.1. Determination of frequency spectra, scaling and implementation of the PCA is the same as for the SOM. Finally, the six input variables for each of the 5200 signals is prepared for CA.

10.2. A FEASIBILITY STUDY BASED ON THE PROPOSED APPROACH

What is the most appropriate distance measure and cluster algorithm for describing the problem? Euclidian distance gives the distance of the "unknown" point from the group mean point and it is usually used for distance calculation. Also, the SOM calculates the Euclidian distance in its algorithm. For ease of comparison between CA and SOM, the Euclidian distance is chosen in this work as a distance measure. Since Euclidian distance does not give a statistical measurement of how well the unknown matches the reference set, and it measures only a

relative distance from the mean point in the group and does not take into account the distribution of the points in the group, another distance, Mahalanobis distance, is applied to PD data in this thesis.

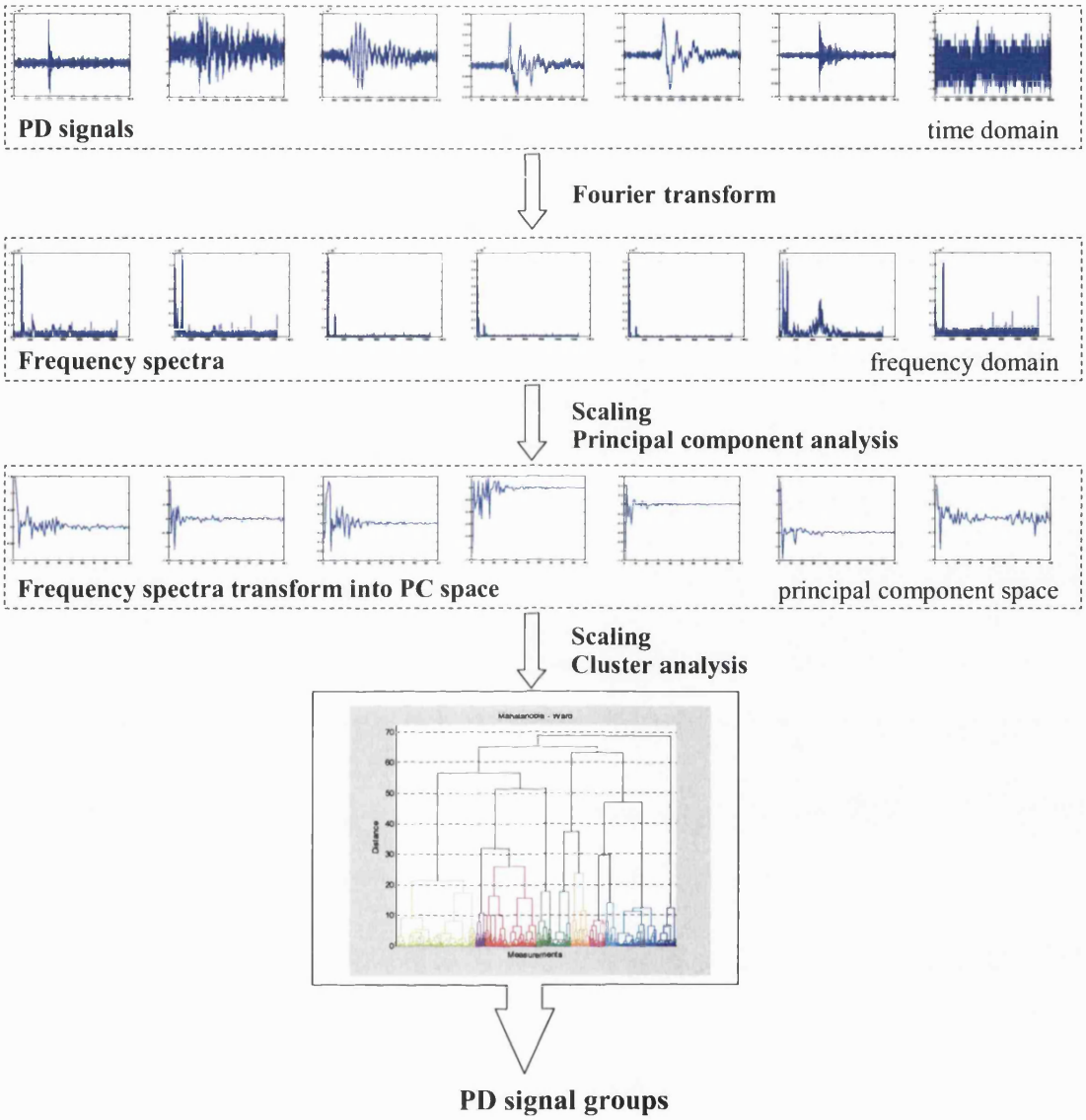


Figure 10.1: Various stages of obtaining groups by CA

Figure 10.2 shows a comparison between Mahalanobis and Euclidian distance and assume that there is a reference set consisting of observations on variables X1 and X2. In this figure, the elliptical shape refers to the Mahalanobis boundary and the circular shape refers to the Euclidian boundary of the reference set. Now assume that two unknown sample points A and B have been added to the system as shown in the figure. By the Euclidian distance method, sample B is likely to be classified as belonging to the group containing sample A because the relative distance from the centre of the circular boundary of these points is the same. However, sample A clearly lies along the elongated axis of the reference group points, indicating that sample A

has much closer likeness to the reference group and is little similar to sample B. This aspect will become clear if one computes Mahalanobis distance instead of Euclidian distance, because the latter method does not take into account the correlation between variables i.e., distribution of points.

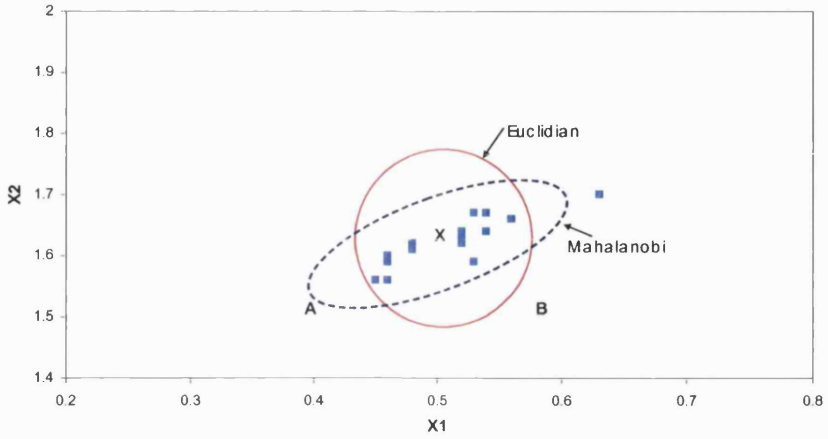


Figure 10.2: Mahalanobis distance and Euclidian distance

Amongst different cluster algorithms, Ward and Average linkage algorithms are implemented herein since single linkage and complete linkage algorithms can cause problems such as chaining or "globular-shaped" clusters, as described in Section 8.9.7. Thus four cases as given in Table 10.1 are going to be considered.

Table 10.1: Four cases with different similarity measures and linkage algorithms.

	Similarity measure	Linkage algorithm
Case 1	Euclid	Ward
Case 2	Euclid	Average
Case 3	Mahalanobis	Average
Case 4	Mahalanobis	Ward

10.2.1. Results of CA

Case 1, which is the process of clustering using the Euclidian distance as similarity measure and Ward linkage algorithm, is summarised in the dendrogram given in Figure 10.3. Similar measurements are joined by links whose position in the diagram is determined by the distance between the measurements. The distance between two successive clusters depends on the number of clusters as given in Figure 10.4. With regard to Figure 10.4 and the guidelines given in Section 8.9.8, the optimal number of clusters can be 5, 7, 10, 13 or 23. If the cut off of the dendrogram is made at distance 8, five clusters are obtained, as shown in Figure 10.5. The first group contains 1894 measurements, the second 1026, the third 828, the forth 1109 and the fifth

343 measurements. When analysing the signals in each groups, the fifth group contains the same signals as group 1 in u-matrix, obtained by SOM, shown in Figure 9.5. These signals, presented in detail in Figure 9.6, have the characteristics of PD in oil as explained later on in Chapter 12.

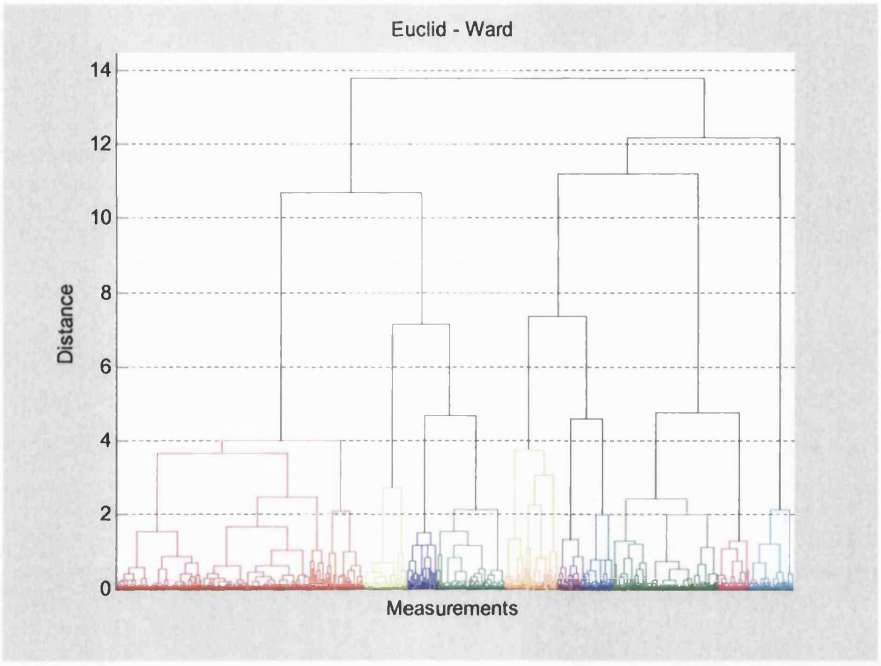


Figure 10.3: Clustering results for Euclid similarity measure and Ward linkage algorithm

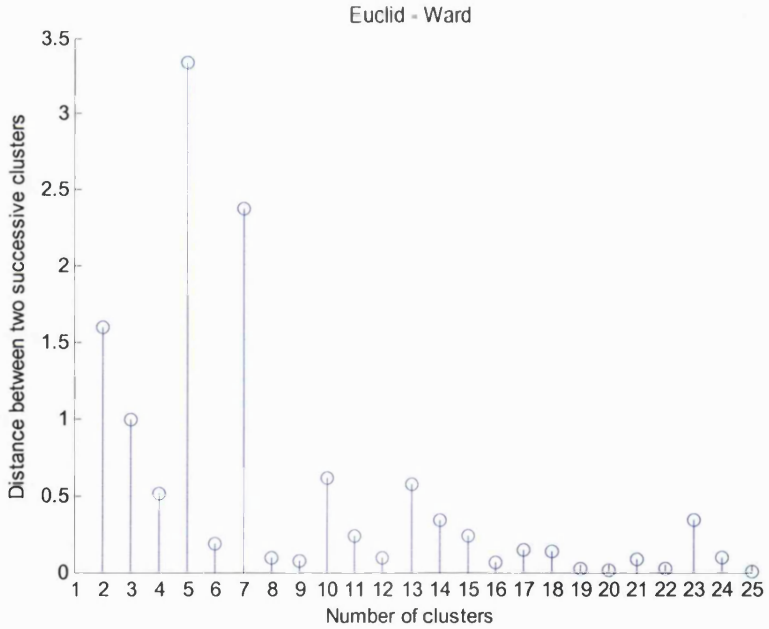


Figure 10.4: Distance between two successive clusters for Euclid similarity measure and Ward linkage algorithm

With only five groups, it is not possible to separate PD signals in oil with small magnitudes. If the dendrogram is cut off at a distance 4.1, PD measurements are grouped into ten clusters, as shown in Figure 10.6.

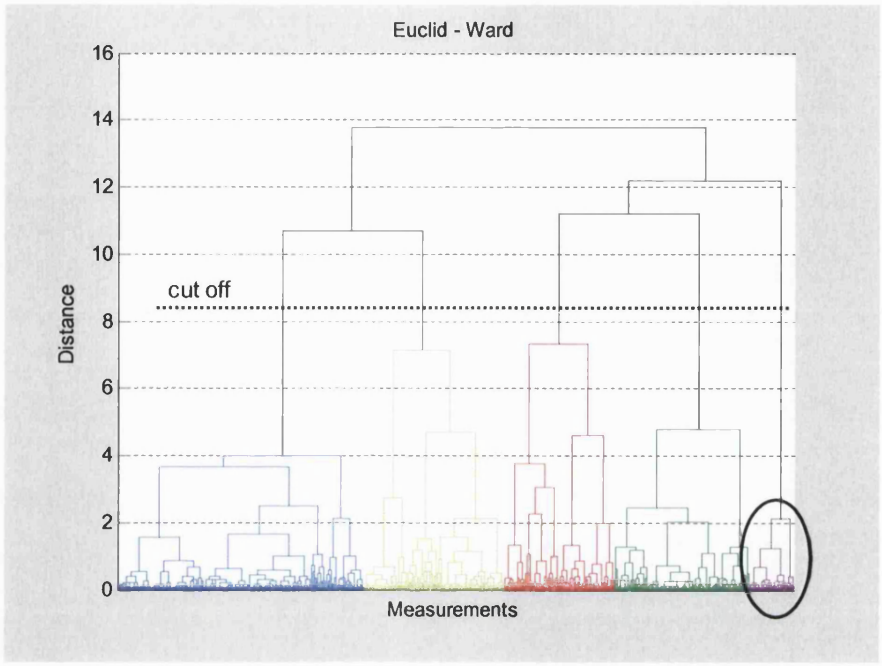


Figure 10.5: PD measurements grouped into five clusters considered Euclid similarity measure and Ward linkage algorithm

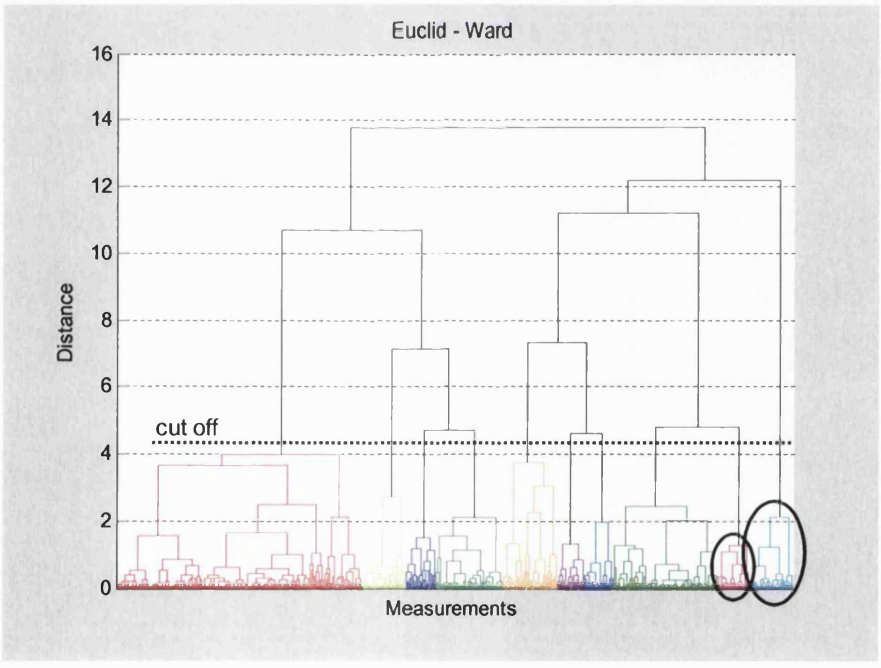


Figure 10.6: PD measurements grouped into ten clusters considered Euclid similarity measure and Ward linkage algorithm

In this case, the group marked with the ellipse at the extreme right side is equivalent to the group 1 in the u-matrix in Figure 9.5 and the smaller ellipse to its left is equivalent the group 4 in the u-matrix obtained by the SOM. The other eight clusters show the characteristics of air discharges and are similar to groups 2, 3, and 5 to 11 obtained by the SOM. A comparison between obtained clustering by SOM and clustering obtained by hierarchical clustering is given in Chapter 11.

Rousseeuw [102] has suggested a graphical display, the silhouette plot, which can be used to assess how well individual observations are clustered. The silhouette shows which objects lie well within their cluster, and which ones are merely somewhere in between clusters.

The silhouette width of observation i is defined as:

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{10.1}$$

where *a(i)* is the average distance between measurement i to all other measurements in the same cluster, and *b(i)* is the minimum average distance of measurement i to measurements in other clusters.

It follows from the formula that $-1 \leq S(i) \leq 1$. If silhouette value is close to 1, it means that measurement is "well-clustered" and it was assigned to a very appropriate cluster. If silhouette value is about zero, it means that that measurement could be assigned to another closest cluster as well, and the sample lies equally far away from both clusters. If silhouette value is close to –1, it means that measurement is "misclassified" and is merely somewhere in between the clusters. The overall average silhouette width for the entire plot is simply the average of *S(i)* for all objects in the whole dataset.

Figure 10.7 shows the silhouette plot of each PD measurement within its cluster. Table 10.2 presents the number and percentage of total number of PD measurements in each cluster as well as cluster average silhouette width. Clusters 6, 8 and 9 contain measurements with negative silhouette width. Since the task of performing CA on PD data is to separate discharges in the transformer and discharges in the air, clusters 5 and 10 are the most important since they have the characteristics of discharges in oil. It can be seen from Figure 10.7 that clusters 5 and 10 have cluster average silhouette width 0.74 and 0.90, respectively. Approximately 10% of all captured data presents discharges in oil, as given in Table 10.2.

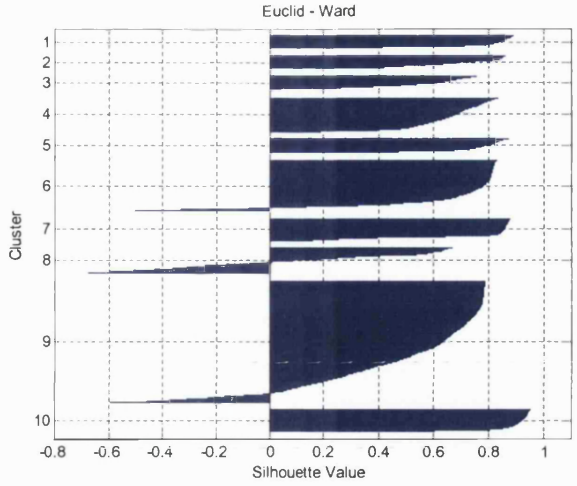


Figure 10.7: Silhouette plot of PD data of 10 clusters for case 1

Table 10.2: Number and the percentage of total amount of PD measurements in each cluster and cluster average silhouette width

Cluster	Number of measurements	Percentage of total amount of PD measurements	Cluster average silhouette width
1	206	3.9	0.78
2	217	4.2	0.68
3	226	4.3	0.51
4	534	10.3	0.66
5	224	4.3	0.74
6	802	15.5	0.64
7	346	6.7	0.76
8	405	7.8	0.09
9	1894	36.4	0.51
10	343	6.6	0.90

The results of clustering procedure for cases 2, 3 and 4 are given in Figures 10.8, 10.10 and 10.12, respectively. Their distances between two successive clusters, depending on the number of clusters, are given in Figures 10.9, 10.11 and 10.13, respectively.

The optimal number of clusters for case 2 (Euclid – average) is 2, 8, 11 and 17, as shown in Figure 10.8. If only two clusters are considered, it can be seen that one cluster consists of 343 measurements belonging to PD in oil, the same as group 1 obtained by SOM, while the other group consists of 4857 measurements. The group of 343 measurements is marked with the vertical ellipse on the right side in Figure 10.8. With the Average linkage algorithm, it is not possible to get the cluster of PDs in oil with smaller magnitudes (horizontal ellipse) as is the case with PDs in group 4 obtained by SOM, although the number of clusters is 17.

The optimal number of clusters for case 3 (Mahalanobis – average) is 6, 9, 13 and 16, as shown in Figure 10.10. If 13 clusters are considered, it can be seen that one group consists of

343 measurements belonging to PD in oil, the same as group 1 obtained by SOM, and is the vertical ellipse on the right side of the dendrogram. Again in this case, with the Average linkage algorithm it is not possible to get the cluster of PDs in oil with smaller magnitudes, as is the case with PDs in group 4 in SOM. They are included in the cluster containing 1053 measurements and are marked with a horizontal ellipse.

The optimal number of clusters for case 4 (Mahalanobis – Ward) is 7, 10, 13 and 19, as shown in Figure 10.12. If 10 clusters are considered, the cut off dendrogram is at distance 28; one cluster consists of 343 measurements belonging to PD in oil, and is the same as group 1 obtained by SOM; this is shown as the vertical ellipse on the right side of the dendrogram. One of the clusters contains 290 signals belong to PD signals in oil with smaller magnitudes as PDs in group 4 in SOM. In Figure 10.12, they are denoted by the smaller ellipse. The other eight groups belong to PDs in the air.

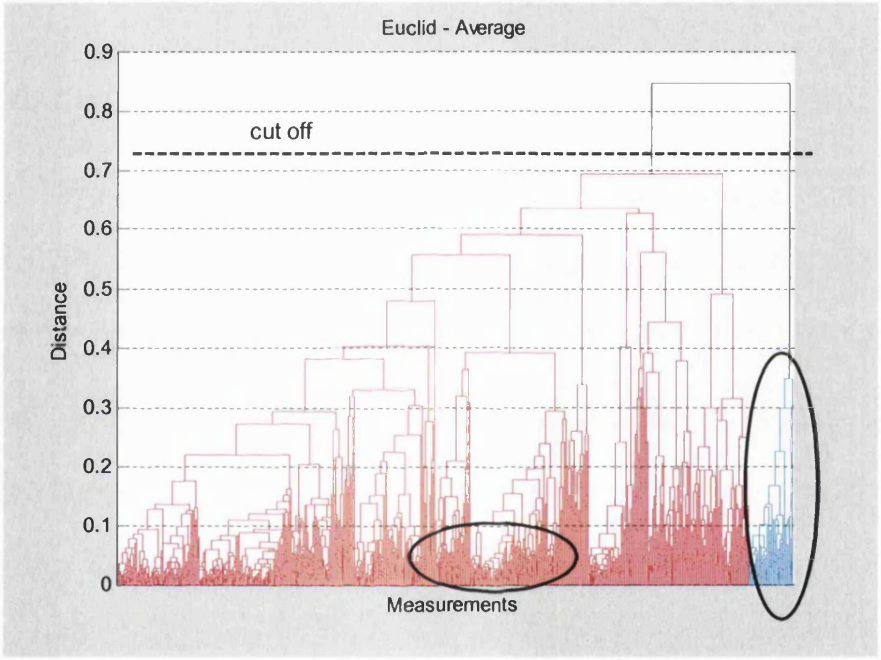


Figure 10.8: Clustering results for Euclid similarity measure and Average linkage algorithm

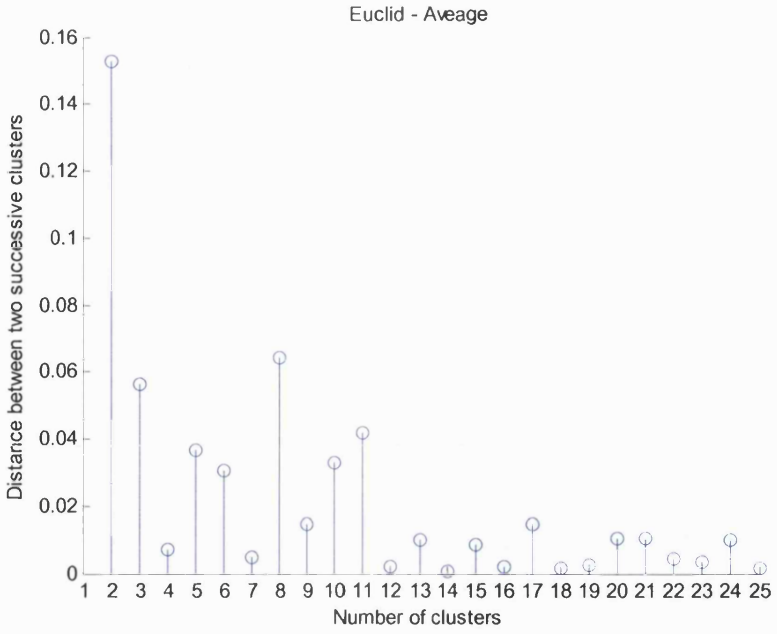


Figure 10.9: Distance between two successive clusters for Euclid similarity measure and Average linkage algorithm

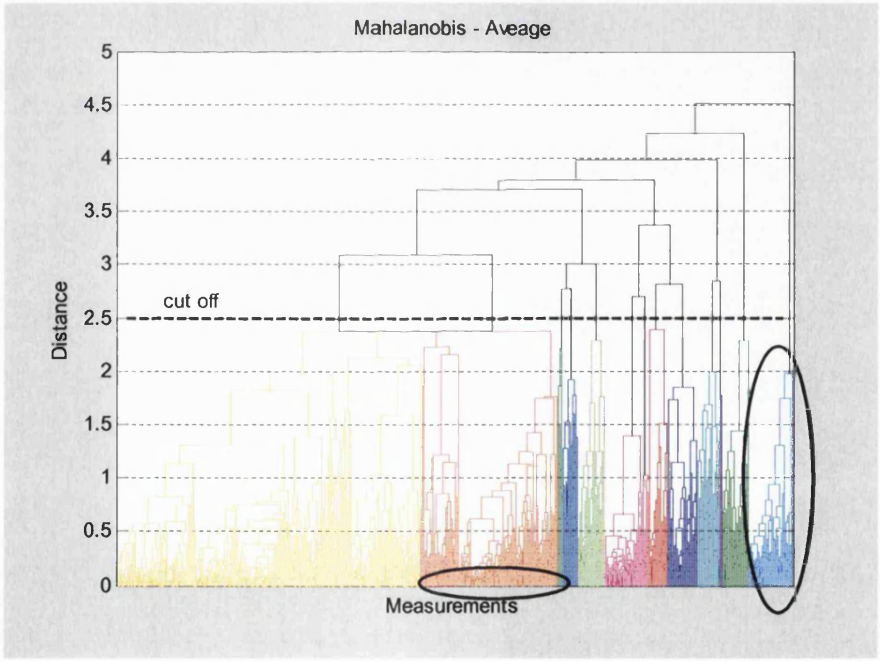


Figure 10.10: Clustering results for Mahalanobis similarity measure and Average linkage algorithm

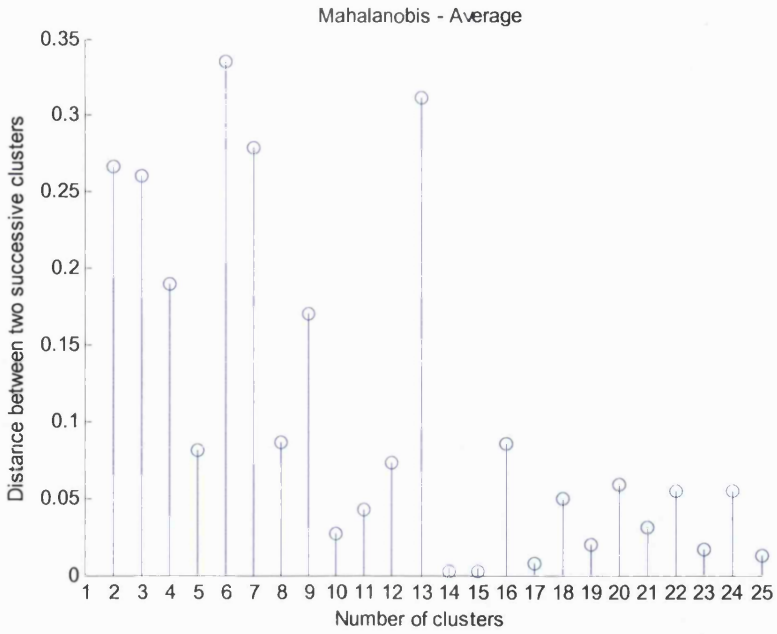


Figure 10.11: Distance between two successive clusters for Mahalanobis similarity measure and Average linkage algorithm

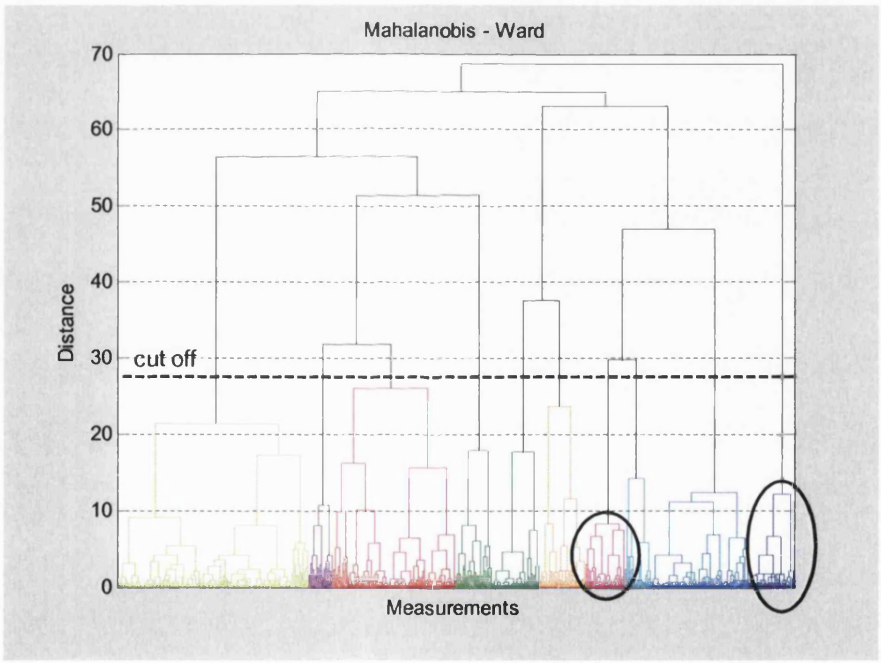


Figure 10.12: Clustering results for Mahalanobis similarity measure and Ward linkage algorithm

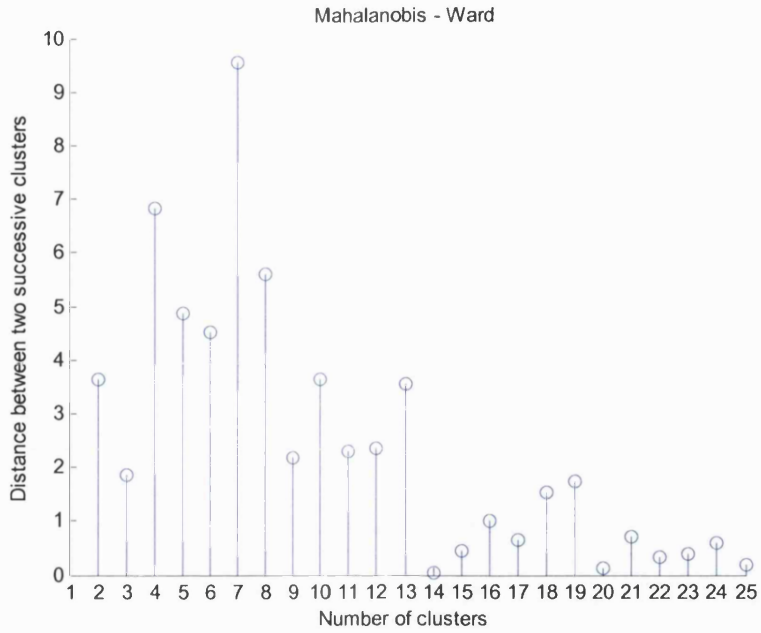


Figure 10.13: Distance between two successive clusters for Mahalanobis similarity measure and Ward linkage algorithm

Figures 10.14 to 10.16 show the silhouette plot for cases 2, 3 and 4, respectively. The cluster containing 343 PD measurements belonging to discharges in oil has always an average silhouette width higher than 0.94. Some of the other clusters contain measurements with negative silhouette width; therefore, they should be moved to the other cluster. Since the focus of this thesis is not in discharges in the air, there is no further investigation as to which cluster measurements with negative silhouette width should be moved.

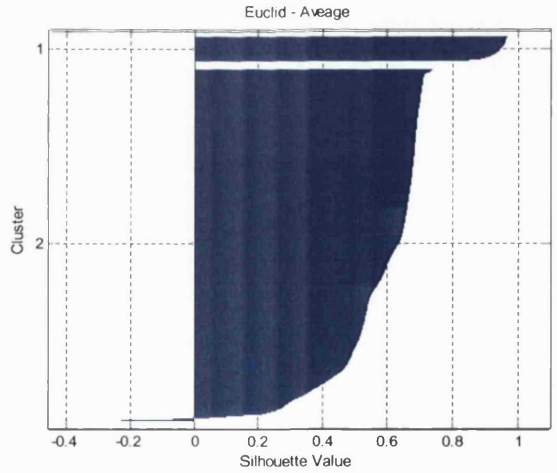


Figure 10.14: Silhouette plot of PD data for case 2

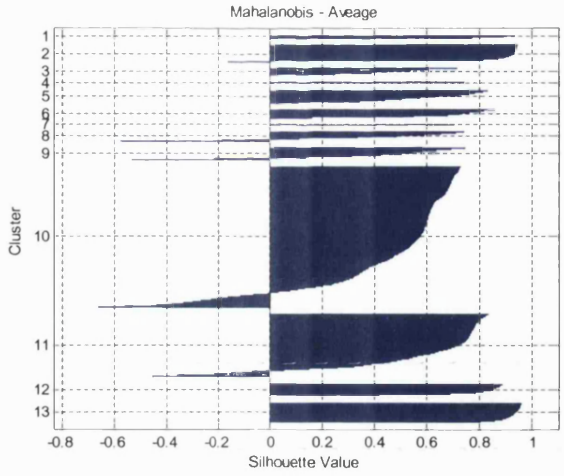


Figure 10.15: Silhouette plot of PD data for case 3

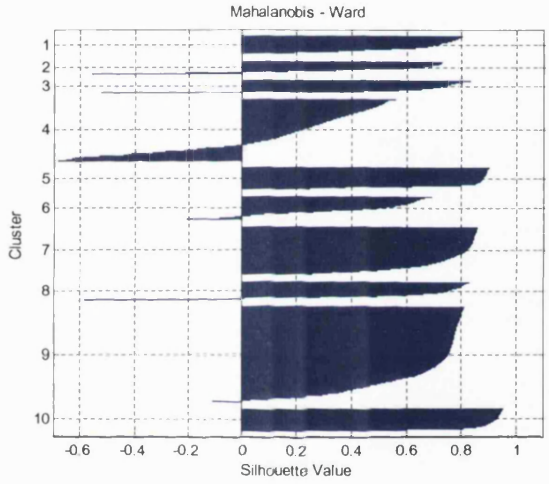


Figure 10.16: Silhouette plot of PD data for case 4

10.3. SUMMARY ON DATA MINING ON THE PD DATA USING THE CA

When applying the six input variables of PD data preprocessed by PCA to CA, the following can be concluded:

- The type of cluster algorithm is a very relevant factor for clustering PD data.
- The Average linkage algorithm caused single samples to join a larger cluster each time, so that there was no true hierarchical structure. Consequently, with this method, it is only possible to separate the oil discharges of high magnitudes. Thus the Average linkage algorithm is not suitable for clustering of captured PD data.
- Ward's linkage method is based on a sum of squares approach and tends to create clusters of similar size. It is the only method to rely on analysis of variance; therefore it

tends to produce clearly defined clusters. Ward's method provides distinct groups with a relatively large membership within each group. Thus the best clustering results were achieved with the Ward linkage method.

- The choice between the Euclidian distance and Mahalanobis distance does not lead to big differences in clustering PD data.
- It was found that only Ward's method continued to perform well by providing distinct groups with a relatively large membership within each group.

11. A COMPARISON OF RESULTS OBTAINED BY SOM AND CA

This chapter deals with a comparison of grouping/clustering results obtained by SOM (Chapter 9) and CA (Chapter 10). All 5200 PDs are considered. The basis for comparison is the u-matrix obtained by SOM.

Figure 11.1 shows the u-matrix obtained by SOM again, as illustrated in Figure 9.4. In addition, 5200 PD signals mapping into eleven groups are shown. In Sections 9.2 and 9.3 it has been demonstrated that two groups contain signals originating from the transformer, as will be shown later in Chapter 12. These two groups are represented by black (group 1) and blue (group 4) dots in Figure 11.1. Since the goal of this work is focused on separation of data originating from the transformer and data originating from transformer surrounding, in the following main attention is given to black and blue dots.

Figure 11.2 illustrates the formation of ten clusters in the u-matrix obtained by CA employing Mahalanobis distance and Ward linkage algorithm, as described in Section 10.2. With reference to Section 10.3 which clearly demonstrates that, when applying the CA techniques for grouping PD data, the best result is achieved by the Ward's linkage algorithm. Therefore groups obtained by Ward's linkage algorithm are transformed into the u-matrix herein. The intention in this figure is also given to groups represented by black and blue dots.

Figure 11.3 presents the formation of ten clusters in the u-matrix obtained by CA employing Euclidian distance and Ward linkage algorithm, as described in section 10.2. Again, the intention is given to two groups, represented by black and cyan dots, respectively.

As can be seen from all the three figures, the group with black dots in the lower right-hand corner is the same. This group consists of 343 measurements belonging to signals of PD in oil with high magnitudes. Since groups represent by blue dots in Figures 11.1 and 11.2 and group represents by cyan dots in Figure 11.3 slightly differ, they are investigated in detail herein.

In order to obtain a better picture in the agreement of groups obtained by SOM and CA, a comparison is given in Tables 11.1 to 11.3. The bold number in the tables represents how many PD measurements in one group obtained with one method correspond to the groups obtained with the other method. For example, in Table 11.1, group 1 obtained by SOM totally (in all 343 measurements) corresponds to group 10 obtained by CA (Mahalanobis – Ward).

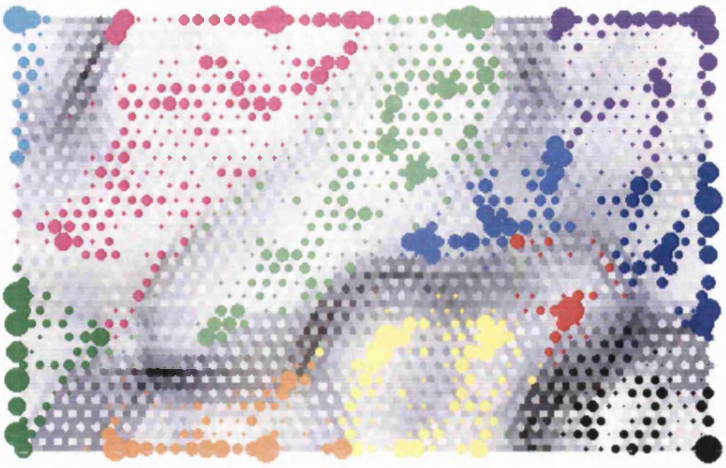


Figure 11.1: Input data transfer into groups obtained by SOM

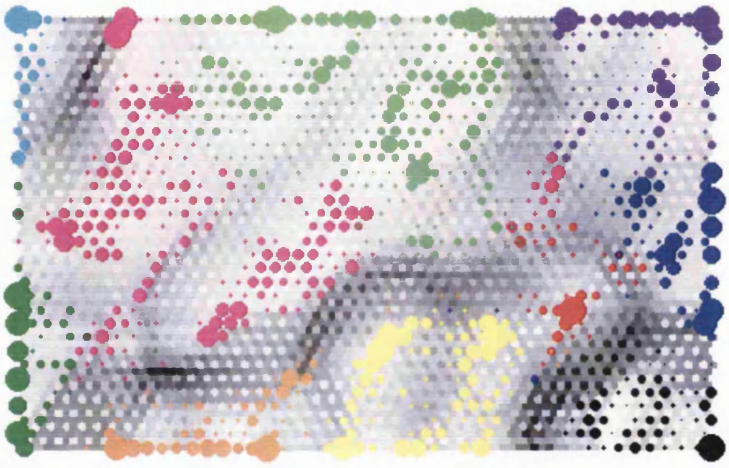


Figure 11.2: Input data transfer into groups obtained by Mahalanobis distance measure and Ward linkage algorithm

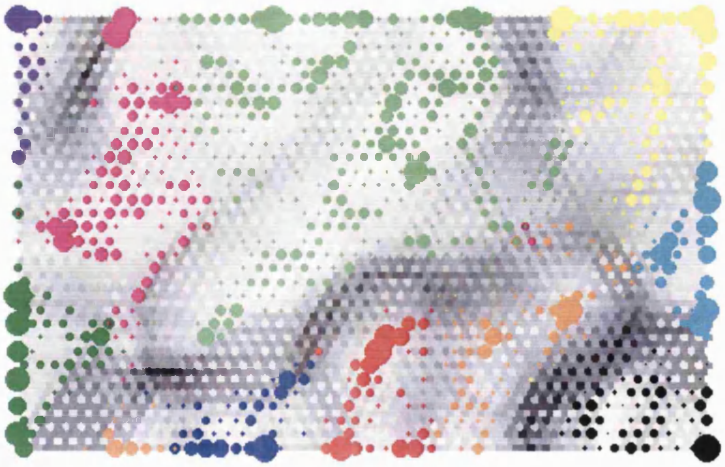


Figure 11.3: Input data transfer into groups obtained by Euclidian distance measure and Ward linkage algorithm

Table 11.1: Agreement of groups obtained by SOM and CA (Mahalanobis-Ward)

		Groups obtained by SOM										
		1	2	3	4	5	6	7	8	9	10	11
Groups obtained by CA Mahalanobis-Ward	1	0	**5**	**1**	**271**	**10**	0	0	0	**1**	0	**2**
	2	0	**1**	**157**	**16**	0	**5**	**1**	0	0	**3**	**15**
	3	0	0	0	0	**1**	0	**1**	**176**	0	**7**	0
	4	0	0	**1**	0	**3**	**1**	**338**	**23**	0	**530**	**58**
	5	0	0	0	0	0	0	0	0	**349**	0	0
	6	0	**334**	0	0	0	**28**	0	0	0	0	0
	7	0	0	0	**2**	**73**	0	0	0	0	0	**14**
	8	0	5	**3**	0	0	**280**	0	0	0	0	0
	9	0	0	**3**	**1**	0	0	**763**	0	0	**609**	**102**
	10	**343**	0	0	0	0	0	0	0	0	0	0

Table 11.2: Agreement of groups obtained by SOM and CA (Euclid-Ward)

		Groups obtained by SOM										
		1	2	3	4	5	6	7	8	9	10	11
Groups obtained by CA Euclid-Ward	1	0	0	0	0	0	**206**	0	0	0	0	0
	2	0	**189**	0	0	0	**28**	0	0	0	0	0
	3	0	0	0	0	**5**	0	**6**	**199**	0	**13**	**3**
	4	0	0	**4**	0	0	**1**	8	0	0	**518**	**3**
	5	0	0	0	**244**	0	0	0	0	0	0	0
	6	0	7	**2**	**53**	**735**	**3**	0	0	0	0	2
	7	0	0	0	**0**	0	0	0	0	**349**	0	0
	8	0	**149**	**155**	**12**	0	**76**	0	0	**1**	0	**12**
	9	0	0	**4**	**1**	**11**	0	**1089**	0	0	**618**	**171**
	10	**343**	0	0	0	0	0	0	0	0	0	0

Table 11.3: Agreement of groups obtained by Mahalanobis-Ward CA and Euclid-Ward CA

		Groups obtained by CA, Mahalanobis-Ward									
		1	2	3	4	5	6	7	8	9	10
Groups obtained by CA Euclid-Ward	1	0	0	0	0	0	0	0	**206**	0	0
	2	0	0	0	0	0	**214**	0	**3**	0	0
	3	0	0	**184**	**41**	0	0	**1**	0	0	0
	4	0	**10**	0	**523**	0	0	0	0	**1**	0
	5	**220**	**4**	0	0	0	0	0	0	0	0
	6	**69**	0	0	0	0	0	**727**	**6**	0	0
	7	0	0	0	0	**349**	0	0	0	0	0
	8	**1**	**183**	0	0	0	**148**	0	**73**	0	0
	9	0	**1**	**1**	**390**	0	0	**25**	0	**1477**	0
	10	0	0	0	0	0	0	0	0	0	**343**

Group 4 obtained by SOM corresponds to group 1 obtained by CA (Mahalanobis – Ward) in 271 measurements, while 16 measurements belong to group 2, 2 measurements to group 7 and 1 measurement to group 9. Since group 4 obtained by SOM, group 1 obtained by CA (Mahalanobis – Ward) and group 5 obtained by CA (Euclid – Ward) contain signals showing the characteristics of PD in oil, they are investigated in detail herein. Other groups containing signals having the characteristics of air discharge are not considered, since the focus of this work is on separation of oil and air discharges.

Data belonging to group 4 obtained by SOM is analysed. Considering the average magnitude of frequency spectra of PD signals for 100 MHz windows given in Figure 11.4, one can conclude that signals can be divided into three subgroups. When analysing these three groups in detail, the following can be concluded. Red dots in the u-matrix in Figure 11.5 represent noise signals while green and blue dots represent PDs in oil (see frequency spectra in Figures 11.6-11.8). Considering the average magnitude of frequency spectra of PD signals for 100 MHz windows given in Figure 11.4, measurements from 140 to 290 transferred into the u-matrix are blue dots. A typical representative PD signal and its frequency spectrum is depicted in Figure 11.6. Measurements 68 to 96, 109 to 119 and 122 to 139 are red dots in the u-matrix and a typical signal and its frequency spectrum is given in Figure 11.7. All other signals are within the green dots. A representative PD signal and its spectrum is given in Figure 11.8.

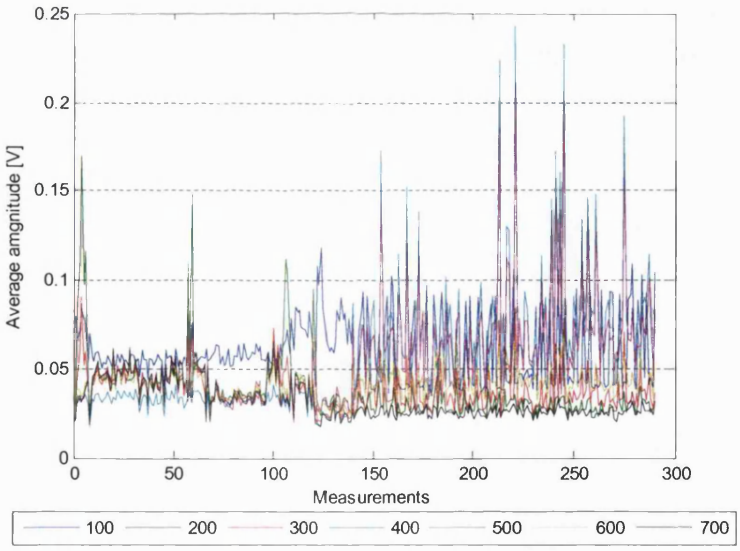


Figure 11.4: Average magnitude of frequency spectra of PD signals for 100 MHz windows of group 4

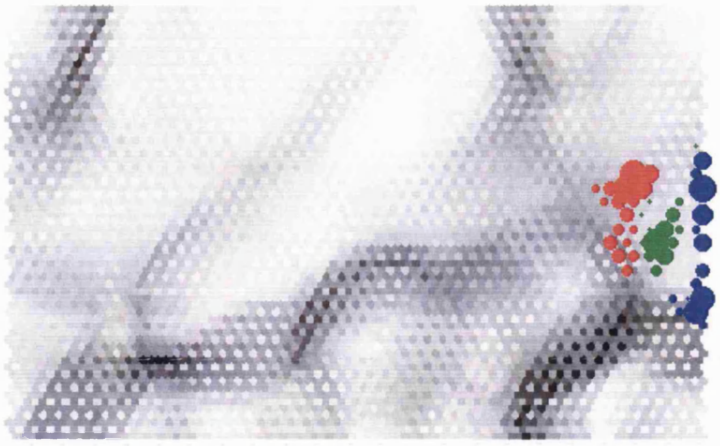


Figure 11.5: Detail consideration of group 4 obtained by SOM

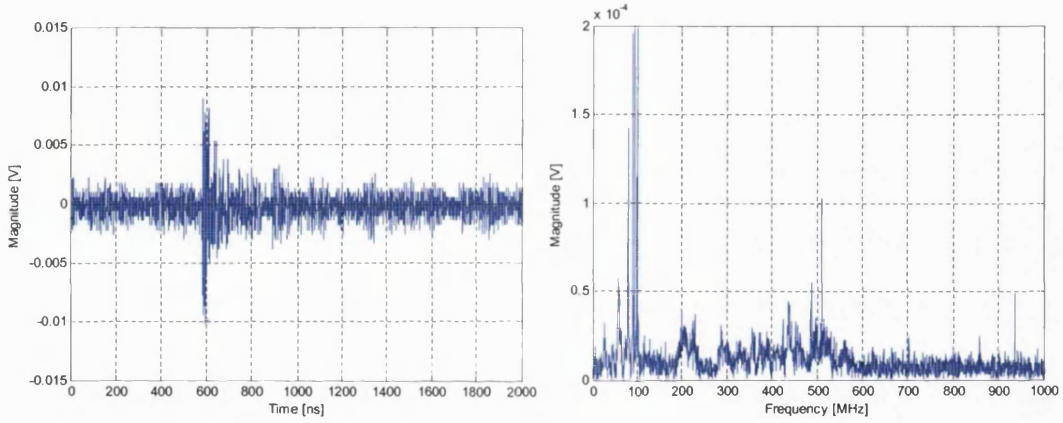


Figure 11.6: Captured signal and its frequency spectrum belonging to blue circles in Figure 11.5

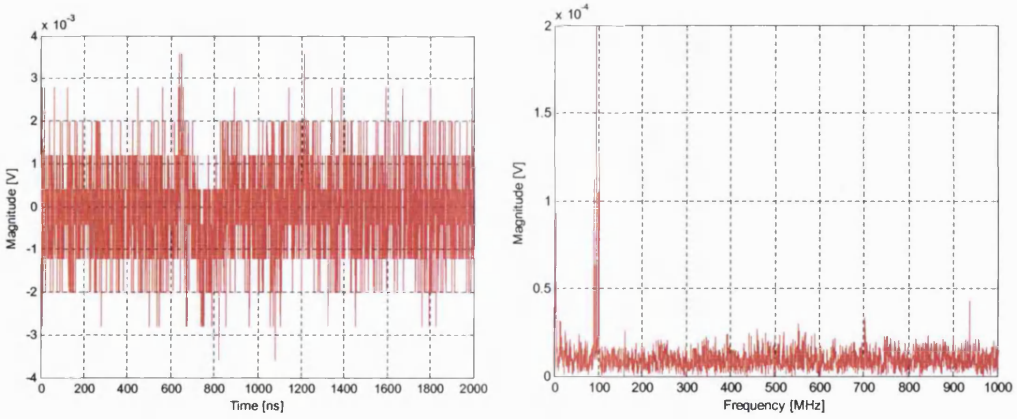


Figure 11.7: Captured signal and its frequency spectrum belonging to red circles in Figure 11.5

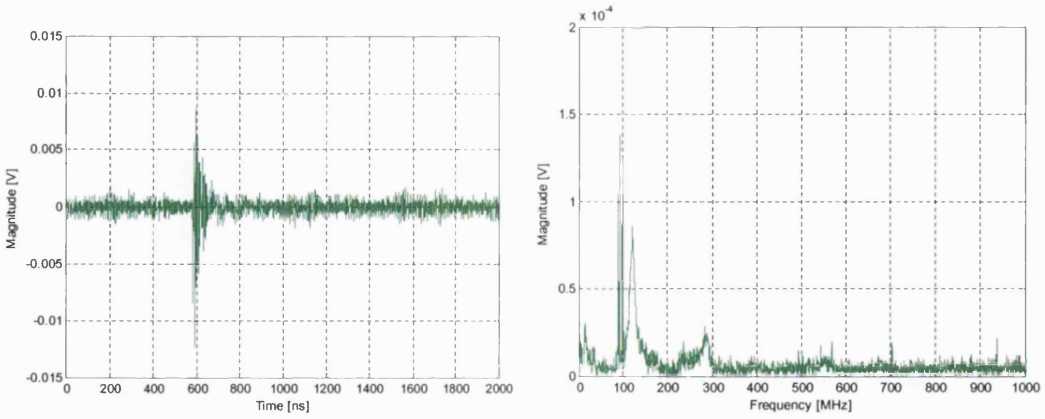


Figure 11.8: Captured signal and its frequency spectrum belonging to green circles in Figure 11.5

A comparison of the group of signals indicating oil discharges of smaller magnitudes obtained by SOM, CA (Euclidian – Ward) and CA (Mahalanobis –Ward), respectively, is demonstrated in Figure 11.9. CA based on Euclidian distance measure and Ward linkage algorithm only gives signals of PDs in oil. However, apart from presenting the signals as obtained by SOM, CA based on Mahalanobis distance measure and Ward linkage algorithm also present signals in the u-matrix belonging to other groups (see also Table 11.1).

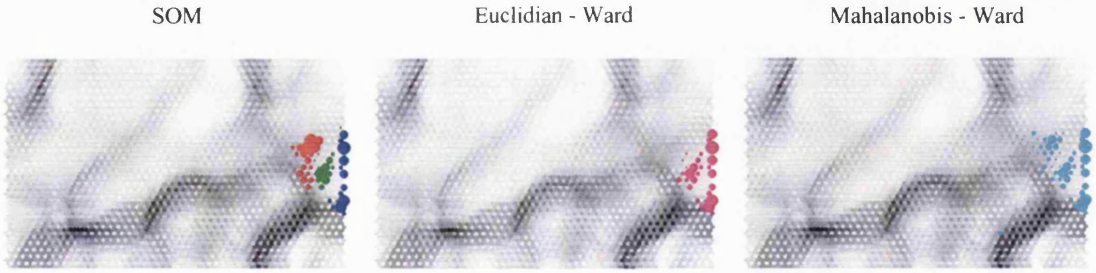


Figure 11.9: Comparison of the group of signals indicating oil discharges smaller magnitude obtained by SOM, CA (Euclidian – Ward) and CA (Mahalanobis –Ward), respectively

With regard to the obtained results, CA based on Euclidian distance and Ward linkage method gives the best results of grouping 5200 measurements captured at the Northfleet substation, and this was so when data is reduced to 6 PCA that serves as input variables for all applied methods.

The strength of SOM and CA is also demonstrated in increased performance when only first three PCs are used as input variables. Decision to choose only three components is based on visualization since measurements can be represented in a 3-dimensional space.

Final results are given herein for Euclidian similarity measure and Ward linkage algorithm. The dendrogram is given in Figure 11.10, while distance between two successive clusters, depending on the number of clusters, is given in Figure 11.11. With regard to Figure 11.11, the optimal number of clusters is 4, 9 or 14. In order to also obtain a cluster of oil discharge of

small magnitudes, 14 clusters are chosen as optimal. Figure 11.12 (3D presentation) and Figure 11.13 (2D presentation) represent 14 groups obtained by CA. The group of discharges in oil with large magnitudes is divided in two groups, denoted by black and blue dots. The group of discharges in oil with small magnitudes is denoted by red dots. Other groups represent discharges in air.

The dendrogram obtained by Mahalanobis similarity measure and Ward linkage algorithm is given in Figure 11.14, while distance between two successive clusters versus the number of clusters is given in Figure 11.15. With reference to Figure 11.15, the optimal number of clusters is 4, 6, 10 or 13. As in the previous case, in order to obtain a cluster of oil discharge of small magnitudes, 13 clusters are chosen as optimal. Figure 11.16 (3D presentation) and Figure 11.17 (2D presentation) represent 13 groups obtained by CA. The group of discharges in oil with large magnitudes is divided in two groups, denoted by black and blue dots. The group of discharges in oil with small magnitudes is denoted by red dots. Other groups represent discharges in air.

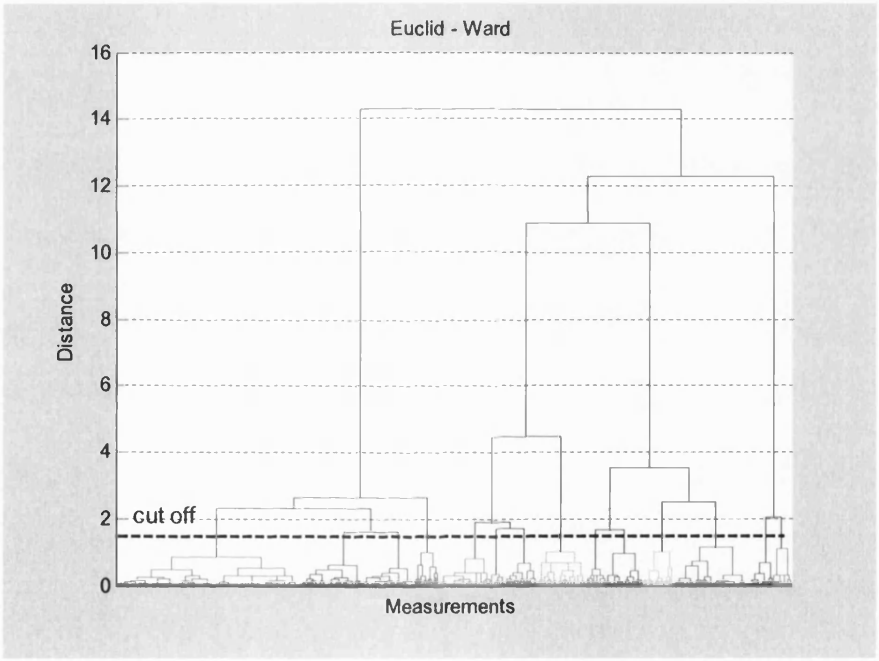


Figure 11.10: Clustering results for Euclidian similarity measure and Ward linkage algorithm, when first three PC are considered

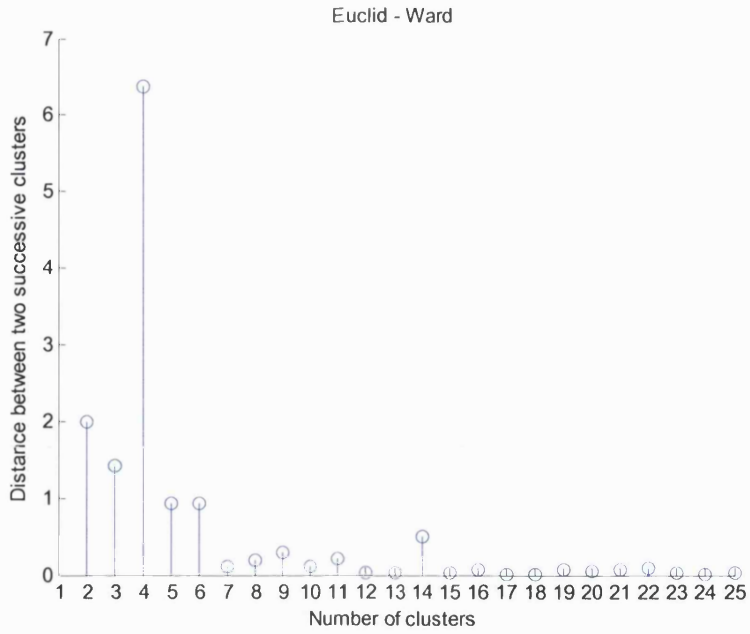


Figure 11.11: Distance between two successive clusters for Euclidian similarity measure and Ward linkage algorithm, when first three PC are considered

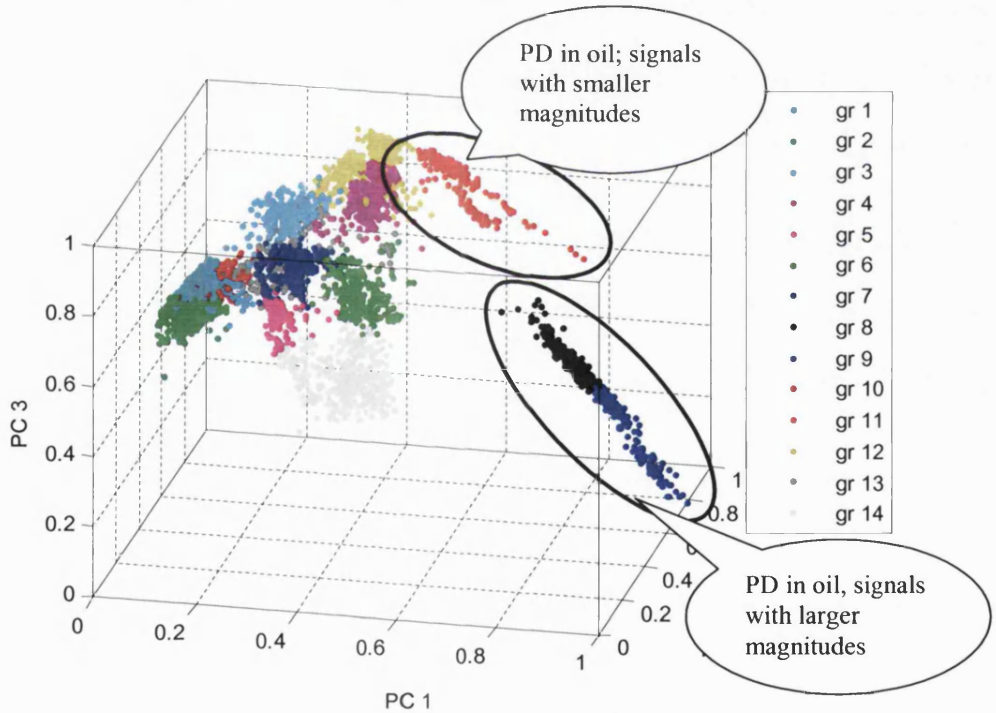


Figure 11.12: 3-dimensional presentation of PD data transferred into PC space when first three PC are considered. Groups are formed for Euclidian similarity measure and Ward linkage algorithm

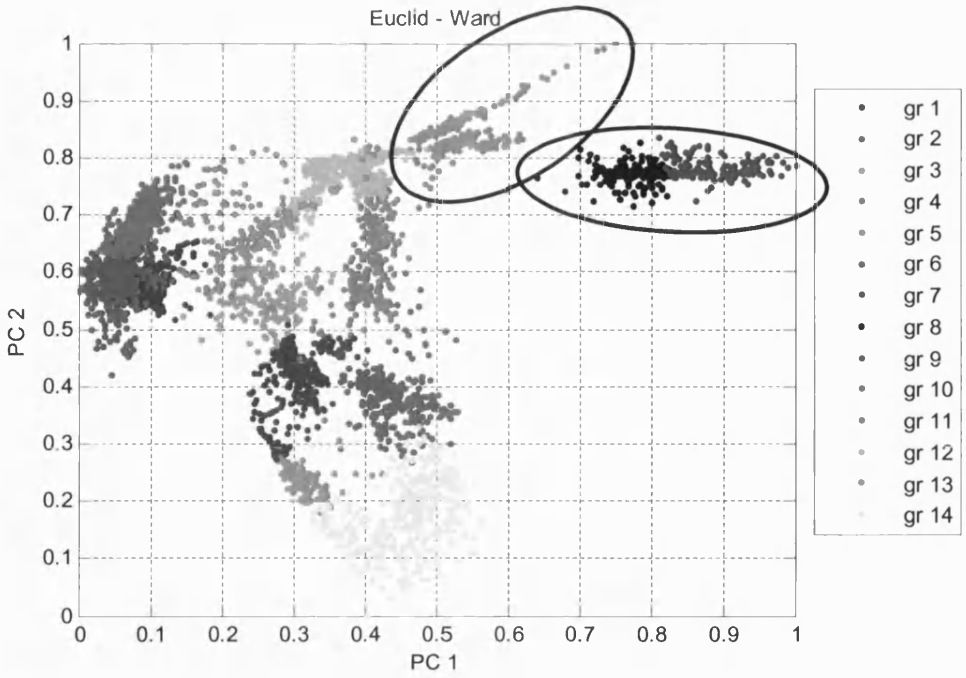


Figure 11.13: 2-dimensional presentation of PD data transferred into PC space when first three PC are considered. Groups are formed for Euclidian similarity measure and Ward linkage algorithm

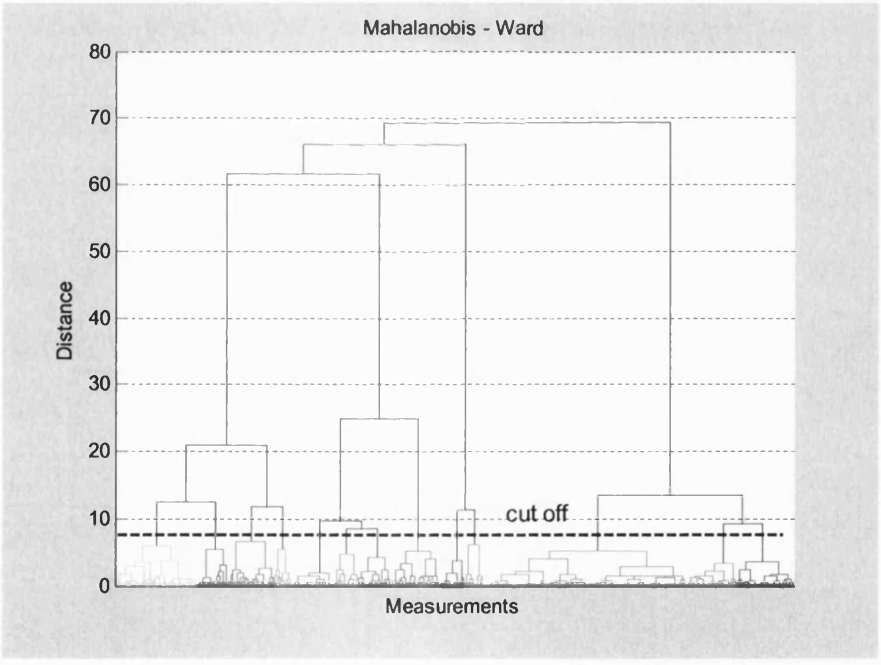


Figure 11.14: Clustering results for Mahalanobis similarity measure and Ward linkage algorithm, when first three PC are considered

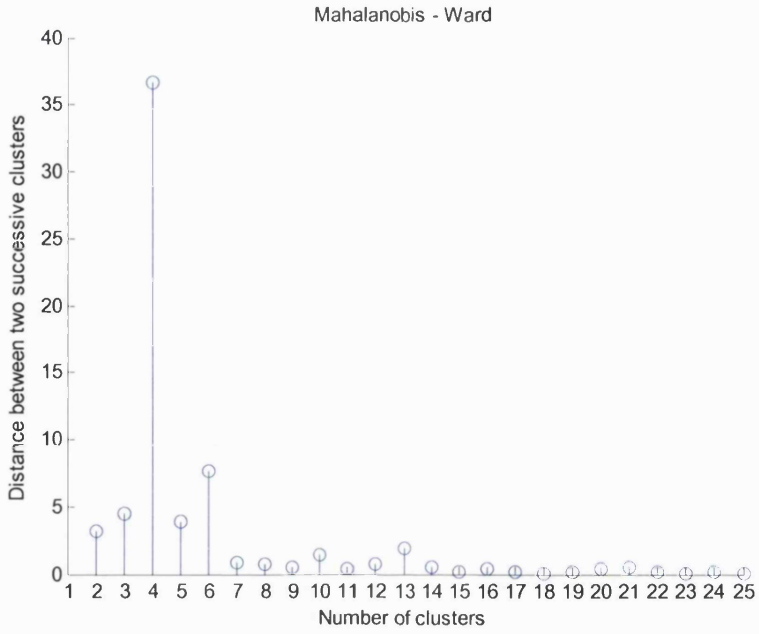


Figure 11.15: Distance between two successive clusters for Mahalanobis similarity measure and Ward linkage algorithm, when first three PC are considered

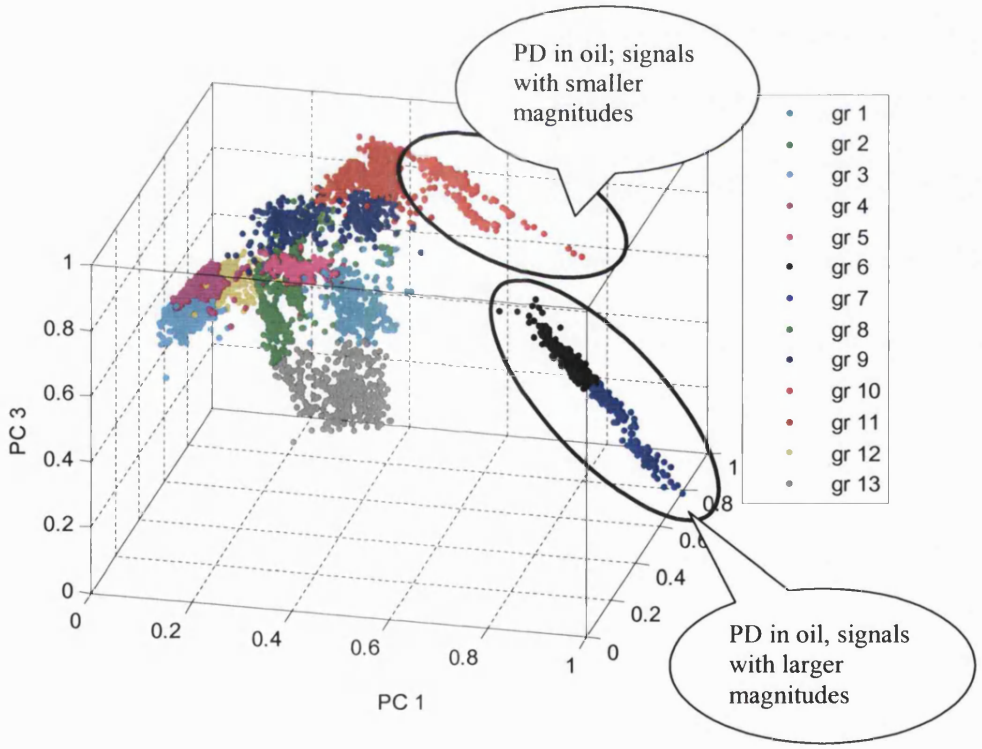


Figure 11.16: 3-dimensional presentation of PD data transferred into PC space when first three PC are considered. Groups are formed for Mahalanobis similarity measure and Ward linkage algorithm

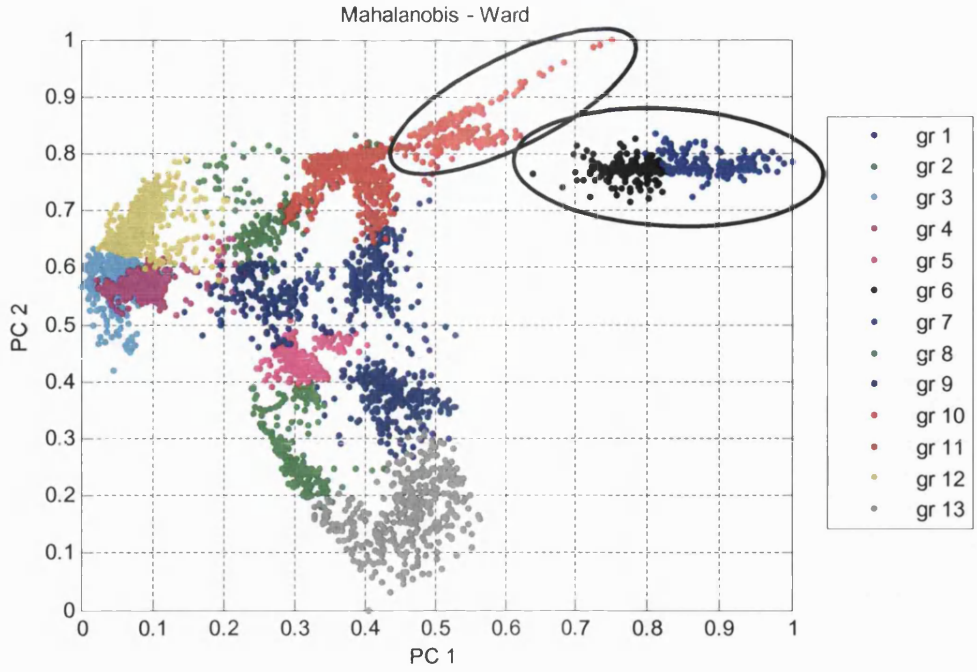


Figure 11.17: 2-dimensional presentation of PD data transferred into PC space when first three PC are considered. Groups are formed for Mahalanobis similarity measure and Ward linkage algorithm

Two important groups of signals, indicating oil discharges are also transformed into the u-matrix obtained by SOM (with 6 PC), as given in Figure 11.18. If a deeper look is indicated in the u-matrix, then this signifies that the group at the lower right-hand corner in the u-matrix can have two sub-groups. With reference to Figure 9.6.a it can be seen that 1st PC varies between value 2 and 4 and 3rd PC between -1 and -2. Additionally, in Figure 9.6.b, showing average magnitude of frequency spectra of PD signals for 100 MHz windows, it can be seen that some signals have higher 500 MHz (purple) and 600 (pale green) window than the others.

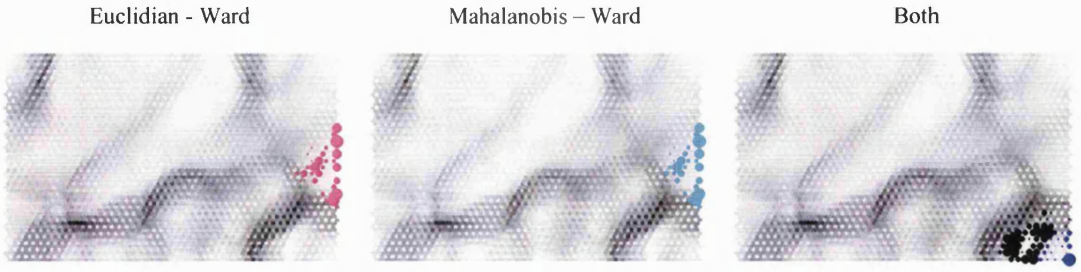


Figure 11.18: Comparison of the group of signals indicating oil discharges smaller magnitude obtained by CA (Euclidian – Ward) and CA (Mahalanobis –Ward) and group of signals indicating oil discharges larger magnitude obtained by both methods

When only first three PCs are applied to the SOM, this can only separate out oil discharges of larger magnitudes (group 1 in Figure 9.5) out of the other signals, and it is not capable of separating oil discharges of smaller magnitude when even the number of training operation is over 208000, as shown in Figure 11.19.

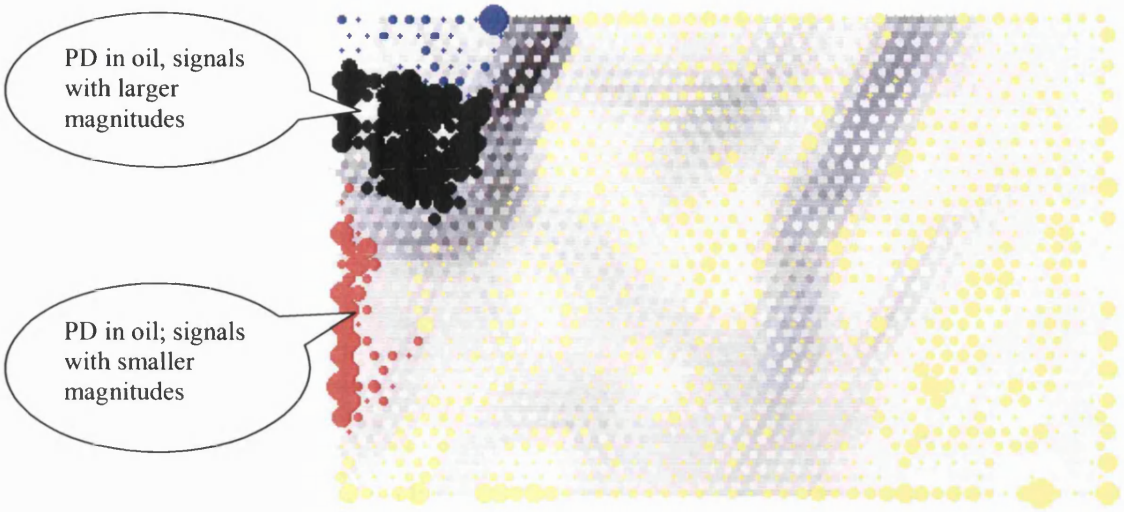


Figure 11.19: U-matrix obtained by SOM, when only first three CA are used as input variables.

11.1. SUMMARY OF OBSERVATIONS

Proposed approaches based on the SOM and CA have been applied for grouping 5200 signals captured on 25th of May 2002 at the Northfleet substation.

By the SOM the final result of group formation can be sometimes subjective. If the final number of groups for classifying signals is not known, one can only visually determine the number of groups based on the u-matrix. The main problem confronted within this work is: where is the real border between two groups? Additionally, selection of the SOM network is based on the empirical obtained formula. Therefore, the appropriate number of learning interactions is obtained by several trials.

On the other hand, the formation of groups obtained by the CA is more evident. The only problem to cope is to find the appropriate number of groups. However, by cutting the dendrogram at the selected threshold value groups are obtained automatically. Moreover, the signal membership to the certain group is exactly determined compared to the SOM, where it is sometimes difficult to determine to which group the signal should belong, especially, when the signal appears at the border between to neighbouring groups. On the other hand, the selection of appropriate CA linkage algorithm, as it is the Ward's in this work, it shows that groups obtained by this algorithm are very coherent as it is not always the case for the SOM.

It can be concluded that SOM gives good results when the final number of groups is known in advance. While when a decision of appropriate number of groups should be taken, easier and better results can be achieved by the CA.

12. RADIO FREQUENCY MEASUREMENT OF DIFFERENT DISCHARGE SOURCES AT THE UMIST LABORATORY

12.1. INTRODUCTION

Investigations have been under way in a number of research laboratories to measure and relate PD signatures to the source of defects. The aim of this chapter tends to justify the statement that there are specific frequency spectra associated with generation of EM pulses from specific PD types. In order to eliminate noisy surrounding, tests were performed in the HV Laboratory at UMIST in Manchester. A practical set-up for generating different discharges to investigate the level of electromagnetic radiation generated by discharges and their detection with wideband helical antenna is presented herein. For the purposes of a comparative analysis, the frequency characteristics are determined separately for the three selected types of PD, i.e. oil discharge, floating discharge and discharge in the air.

12.2. MODELS OF DEFECTS

In order to investigate the likely behaviour of a high voltage equipment with defects, model studies are usually performed. Any experiment with models demands sample preparation and controlled experimental conditions. Various discharge phenomena occurring in a real –life situations were simulated and studied. For the purposes of this study, the following PDs were generated:

Oil discharge, produced with a point/pressboard/plate electrode arrangement with the distance between the point electrode and the pressboard about 5mm, as shown in Figure 12.1. The pressboard is about 6mm thick. The arrangement is immersed in transformer oil during the test. The system is discharge-free up to the test voltage level of 40kV.

Floating discharge, produced with a point/pressboard/ plate electrode arrangement, with two irregular balls made of aluminium foil floating in oil near the HV electrode, as shown in

Figure 12.2. These two aluminium balls are physically half immersed in oil and half floating above the oil. Electrically they are also, a floating potential, which depends on how voltage is induced/coupled on to them. Only under a lower voltage of up to 3~4kV, a discharge can be produced.

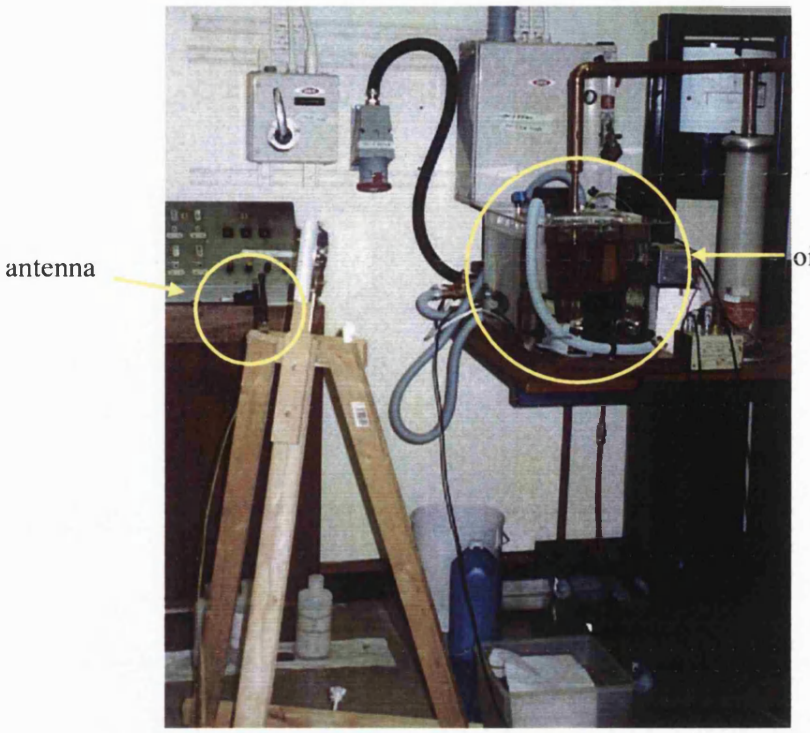


Figure 12.1: Oil discharge source

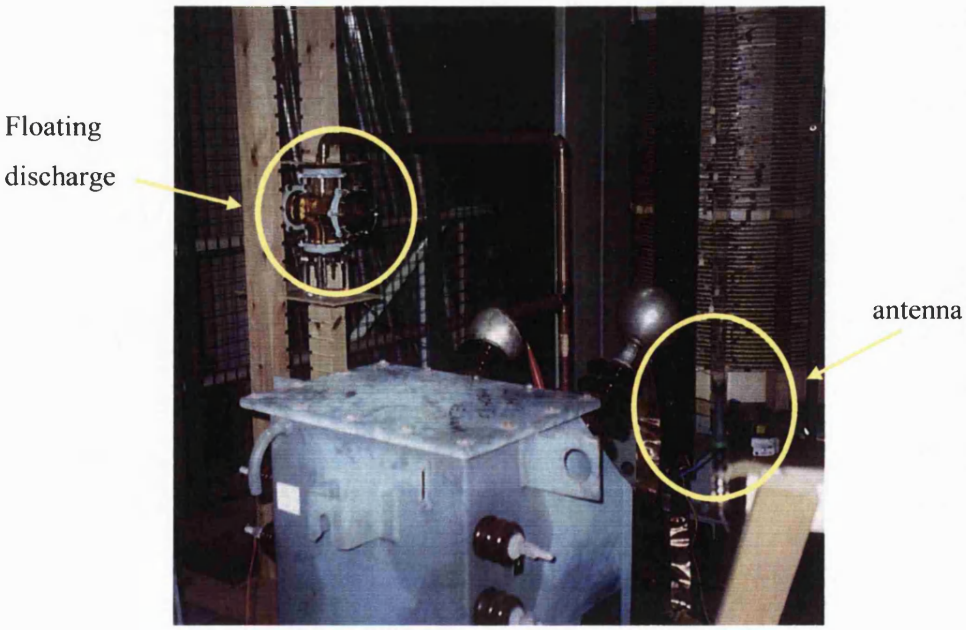


Figure 12.2: Floating discharge source

The 11 kV transformer winding is used for generating *air discharges*, as shown in Figure 12.3. The winding has 72 discs and each disc contains about 6 to 7 turns depending on their physical position. It is a continuous disc type transformer winding. Discharge sources were connected to various disc positions to test how the position of the discharge may affect the detection of the discharge signals.

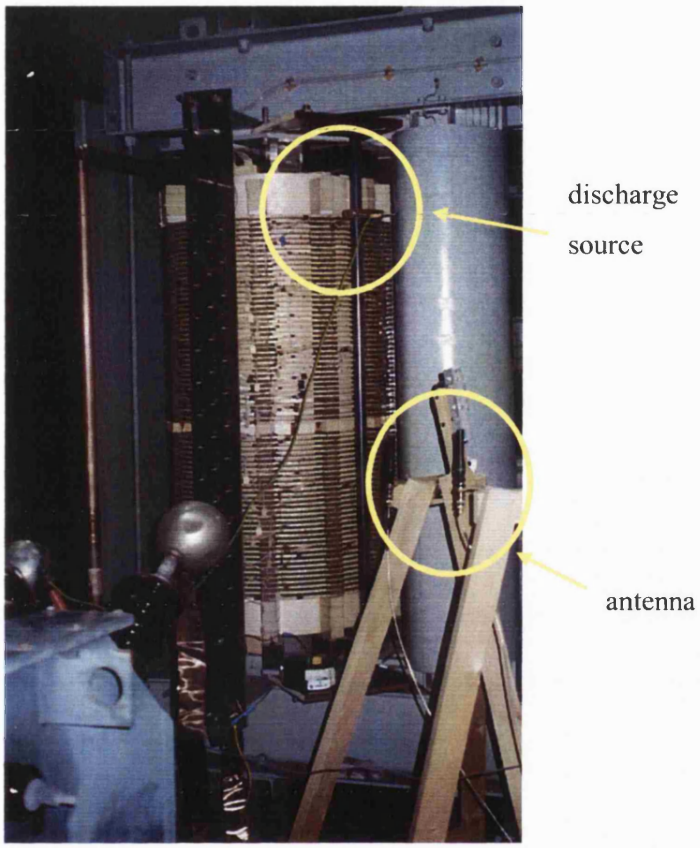


Figure 12.3: Transformer winding used for generating discharge sources

12.3. MEASUREMENT CONFIGURATION

The electromagnetic radiation generated by different types of discharges was measured, similar to the Northfleet substation, by a helical (frequency range 300-500 MHz) antenna, digital phosphor oscilloscope, and laptop. The oscilloscope was set to the "FastFrame" acquisition mode (see Section 7.2), of 100 frames of 5000 samples each. The measurement window was 2μs with $\Delta t = 0.4 \cdot 10^{-9}$ s.

12.4. MEASUREMENT RESULTS AND ANALYSIS

Measurements of discharges can be disturbed by noise originating from communication services and broadcasting stations. Thus, the antenna captured the background noise before

capturing any type of discharges. Figure 12.4 shows the background noise signal and its frequency spectrum. It can bee seen that the frequency spectrum of the background gives signals from FM radio (at 100 MHz), pagers (at 150 MHz) and DAB (at 220 MHz). The magnitude of the noise signal is less than 2 mV.

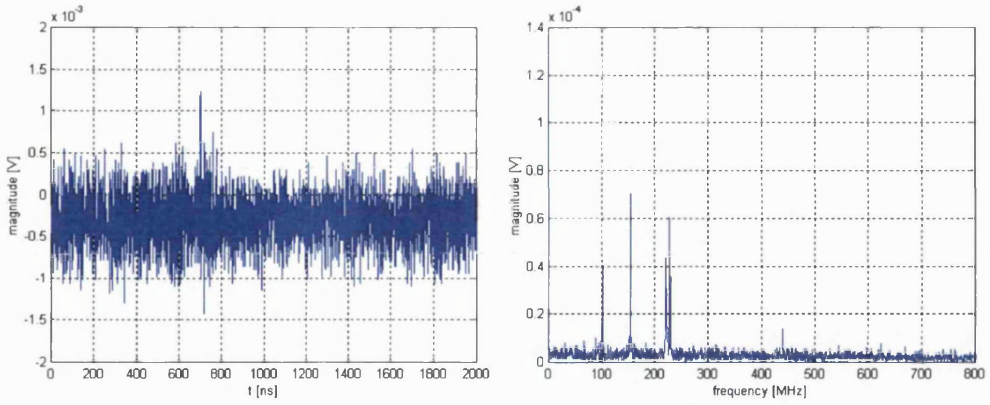


Figure 12.4: The background noise of the signal and its frequency content

Figure 12.5 shows a typical waveform of floating discharge along with frequency spectrum. The duration of the captured signals is around 100 ns. Floating discharge contains high frequencies up to 600 MHz.

The influence of the distance between discharge source and antenna was also studied. The antenna was placed 219 cm, 230 cm and 326 cm, respectively, away from the PD source. Figure 12.6 compares waveforms measured at different locations from the source. As would be expected, the longer the distance the smaller is the magnitude of discharges. Importantly, the frequency spectra of discharges are different in magnitudes but have practically the same dominant frequency components.

Two examples of oil discharges are given in Figures 12.7 and 12.8. Captured signals consist of single pulses (Figure 12.7) or multiple pulses (Figure 12.8). The occurrence of the second discharge was approximately 600 ns after the first double pulse. Duration of pulses is around 200 ns. Single as well as multiple discharges in oil contain dominant frequencies below 300 MHz.

Figure 12.9 typifies an example of air discharge and its frequency spectrum. In general, air discharge contains frequencies below 150 MHz. Duration of air discharges is longer then 800 ns. Generated air discharges were weak compare to discharges in oil or the floating discharges; hence, the magnitudes of air discharges are small compared to magnitudes of discharges in oil or floating discharges. Similar wave shapes of signals were obtained for the disk position at the top and the bottom of transformer winding.

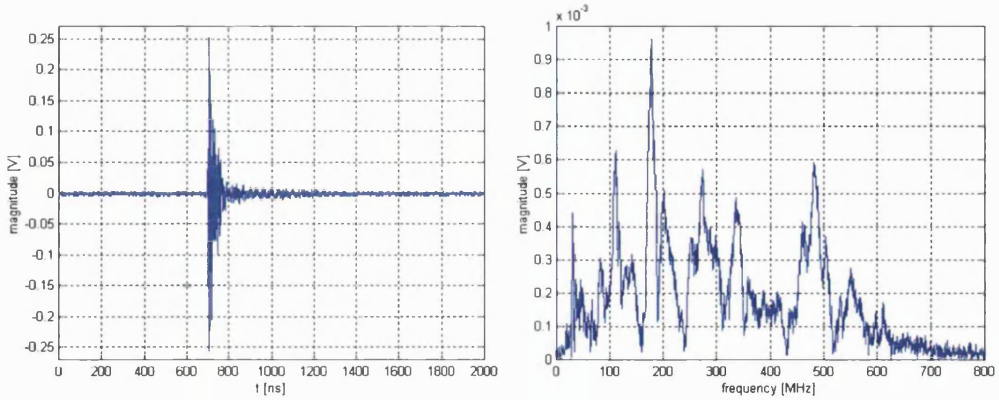


Figure 12.5: PD signal caused by floating discharge and its frequency spectrum

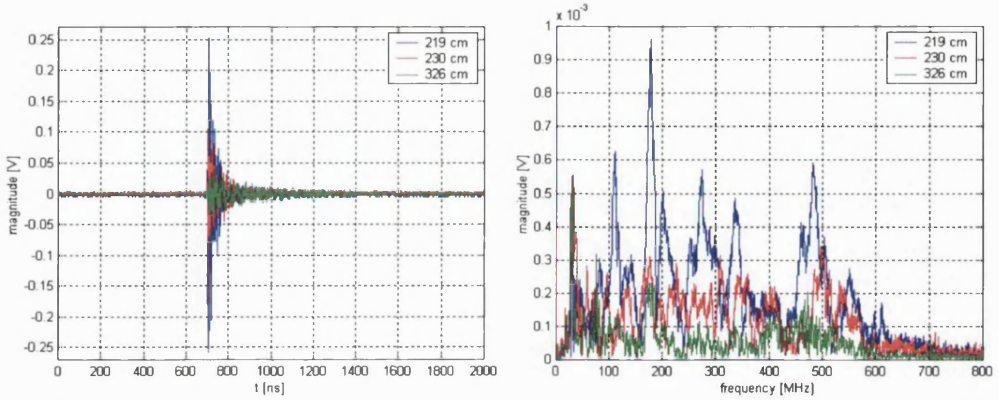


Figure 12.6: PD waveforms measured at different locations and their frequency spectra

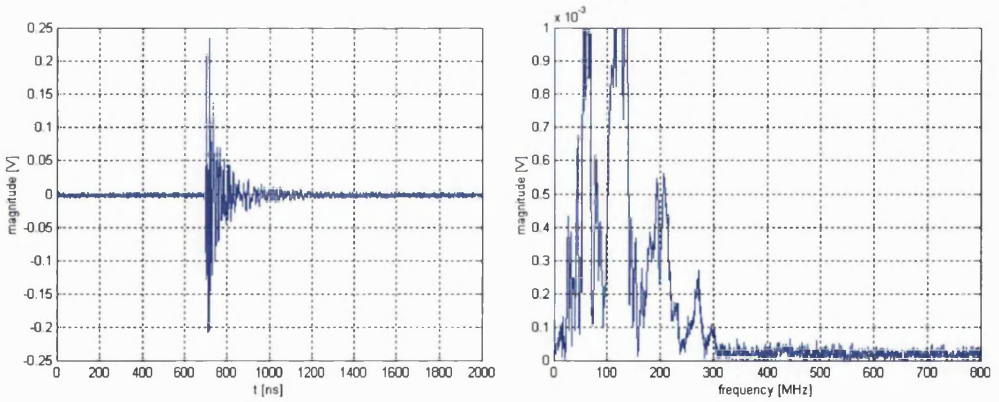


Figure 12.7: Single pulse caused by oil discharge and its frequency spectrum

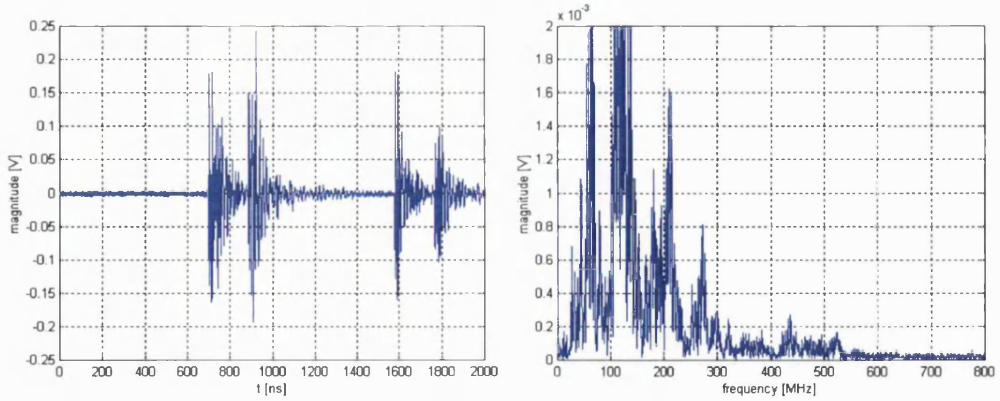


Figure 12.8: Multiple pulses caused by oil discharge and frequency spectrum

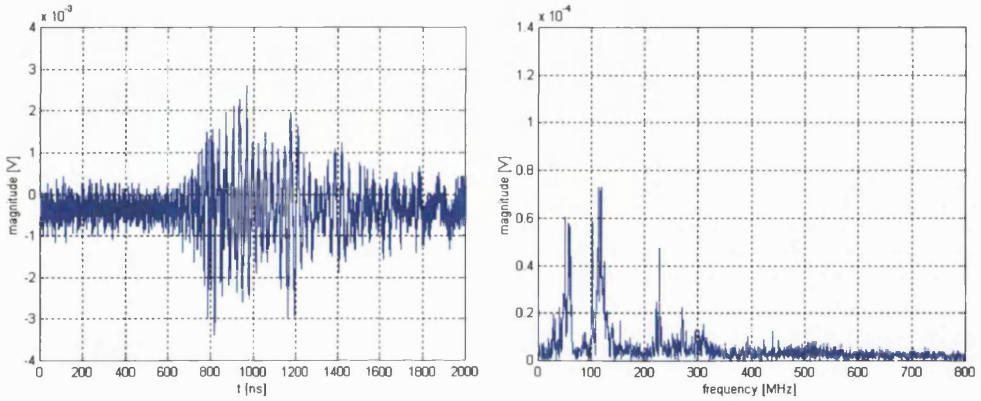


Figure 12.9: The shape of air discharge and its frequency spectrum

12.5. SUMMARY

An oscilloscope equipped with the antenna and connected to a laptop proved to be a useful device in discharge type identification. According to the measurements, it is possible to evaluate and classify discharges by measuring the signals in the time domain and via the frequency spectra of discharges. The difference between wave shapes and dominant frequencies can help with PD classification.

If one compares captured data in the Laboratory and data obtained at the Northfleet substation, one can conclude that discharges in oil contain much higher harmonic components (above 200 MHz) than discharges in the air. This effectively means that tests performed in the Laboratory have confirmed that PD measurements obtained at the Northfleet substation also contain PDs originating from the transformer.

13. METHOD OF MOMENTS APPLIED TO PD DATA

13.1. INTRODUCTION

One of the tasks of this work is focused on dimensionality reduction of captured data. As mentioned before, each measurement consists of 5000 samples. More than 5000 measurements were captured at the substation. Therefore, before applying any of methods for data clustering, the need for feature extraction of signals is essential.

One of the feature extraction methods is the method of moments, which is briefly described in Section 8.7.4. Cavallini et al [104]-[106] have implemented this method for PD signals captured on generators as well as other power system elements. With this approach, it is possible to separate different types of PDs and noise signals. Equations (8.6) and (8.8) enable the presentation of each PD pulse in the time-frequency plane. The main advantage of this method lies in its visualization. Each PD pulse is represented as a point in a two-dimensional space of σ_T and σ_F plane.

Since the method is simple and good for visualization, it was implemented on captured PD signals in this work.

13.2. PD MEASUREMENTS CAPTURED AT THE NORTHFLEET SUBSTATION IN TIME-FREQUENCY PLANE

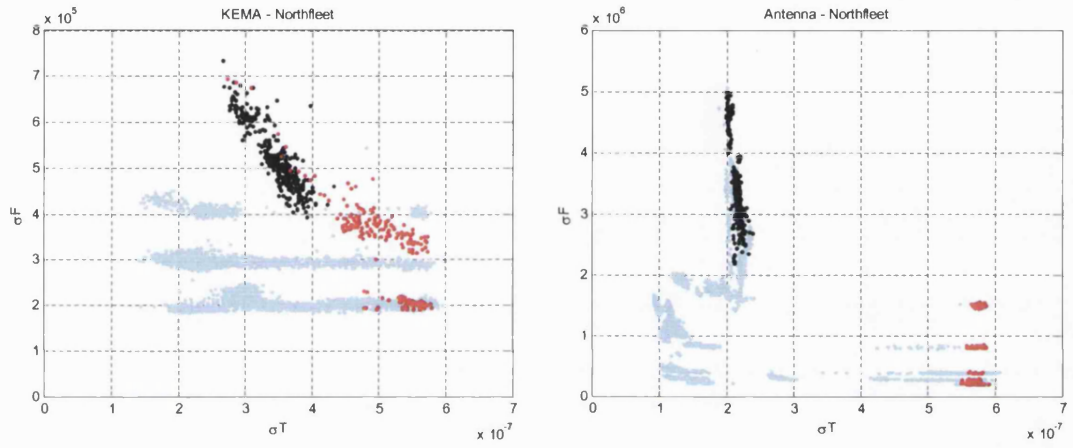
This section deals with σ_T and σ_F results for 5200 measurements captured on the 25th of May 2002 at the Northleet substation by helical antenna and KEMA probe. Figure 13.1 shows the time-frequency plane of all signals captured by the KEMA probe (Figure 13.1.a) and by the helical antenna (Figure 13.1.b). Since the helical antenna was placed at four positions around the transformer (see Figure 7.1), Figure 13.2 represents results for four placements of the helical antenna.

In Figure 13.1, black dots represent PD signals of higher magnitudes originating from oil, while the red dots represent PD signals of smaller magnitudes originating from oil, as obtained by CA using the Euclidian distance measure and Ward linkage algorithm (see Section 10.2). It can be seen that oil PDs of larger magnitudes can be separated from the other discharges, while

oil PDs of smaller magnitudes coincide with discharges in the air. This means that signals of different shapes are mapped on to the same σ_T and σ_F values.

Unfortunately, the noise content and PD signals in this study are not known a priori, and hence when applying this method to the PDs captured during this research, different shapes of signals are mapped together.

In Figure 13.2, H1, H2, H3 and H4 signify different positions of the antenna. Purple dots represent signals captured at position H1, blue colour at position H2, yellow colour at position H3 and green colour at position H4. Signals captured at only position H4 have the characteristics of oil discharges, and therefore it can be concluded that the radiation of PDs is detected by antenna placed at position H4, i.e. radiation of PDs is from the 400 kV, c-phase bushing (see Figure 7.1).



a) b)

Figure 13.1: Time-frequency plane of signals captured at the Northfleet substation by KEMA probe (left) and helical antennas (right)

13.3. Comparison of PD measurements captured at the Northfleet and the Melksham substation in time-frequency plane

Radio frequency measurements were also performed at the Melksham substation on the 400/275 kV, 750 MVA power transformer. As ascertained by the DGA, the transformer does not suffer from any PDs. The aim of the measurements is to compare signals obtained at the Northfleet with the measurements obtained at the Melksham. The KEMA probe was also installed on the oil valve of the transformer. Results of analysis of signals obtained by the KEMA probe and antennas have shown that only discharges in the air were captured, as expected.

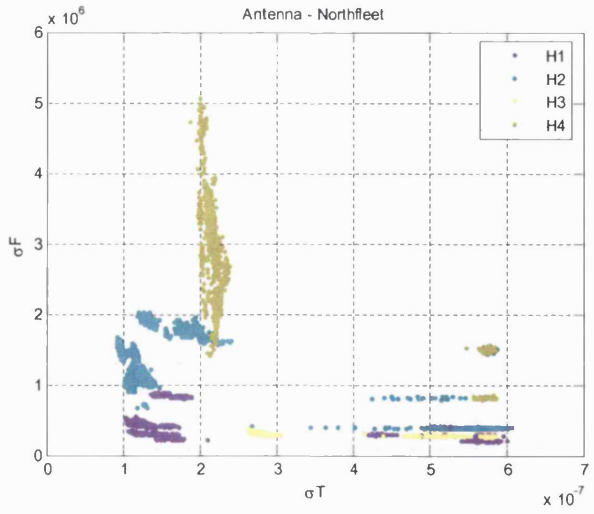
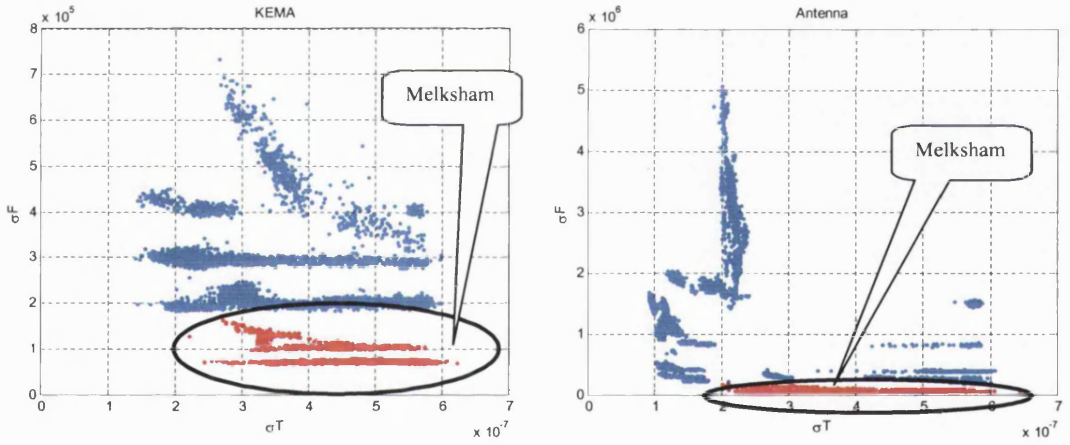


Figure 13.2: Time-frequency presentation of signals captured by helical antennas placed at position H1, H2, H3 and H4 around the transformer

Figure 13.3 shows σ_T and σ_F results of signals captured at the Northfleet substation (grey dots) and at the Melksham substation (red dots), by the KEMA probe (Figure 13.3.a) and by helical antenna Figure 13.3.b).



a) b)

Figure 13.3: Time-frequency presentation of signals captured at Northfleet and Melksham substations

13.4. MEASUREMENTS OBTAINED AT THE UMIST LABORATORY IN TIME-FREQUENCY PLANE

The projection of signals measured at the UMIST Laboratory, into a time-frequency plane was also performed. Results of projection are given in Figure 13.4. Blue dots are the signals originating from floating particles (FP), green dots are signals originating from the oil

discharges, red dots represent background noise (BGN), while pale blue dots represent signals generated by discharges in the air. It is clearly evident from the figure that signals of different sources have different positions in the σ_T-σ_F plane.

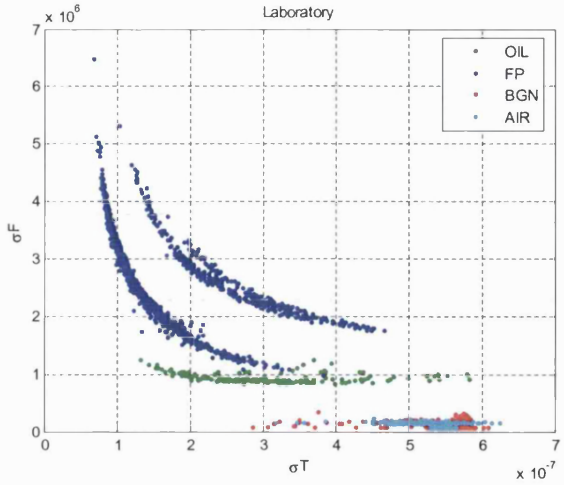


Figure 13.4: Time-frequency presentation of signals captured by the antenna in the Laboratory

13.5. SUMMARY

Feature extraction from signals, based on the method of moments, illustrates that validating the information retained in equations (8.6) and (8.8) is not the best solution for the PD data obtained during this research. In particular, measurements obtained by the KEMA probe at the Northfleet substation show that signals of different shapes can be mapped on to the same σ_T and σ_F values. With regard to Cavallini et al, who first implement this method for PD data, they acquired practical experiences during three years of many tests in a laboratory and the industrial environments. Unfortunately, it was not possible to establish as to how many samples contained their PD signals. One possible explanation as to why different PD shapes map together is that the original signal consists of 5000 samples. The presented feature extraction method transfers the 5000 values into only two parameters. The approximation of the original signal with only two parameters is not sufficient in the presented case.

14. CONCLUSIONS AND FUTURE WORK

14.1. INTRODUCTION

Power transformers are a major power system equipment. Their reliability not only affects the electric energy availability to the supplied area, but also affects the economical operation of a utility. For example, a fault in a distribution transformer may leave thousands of homes without heat and light, and ae fault in a step-up transformer in a power generation plant may cause the shutdown of the attached generation unit.

Under the deregulation policy of electric systems, each utility is trying to cut its costs, and the prevention of accidental loss is much more important than before privatisation. The capital loss of an accidental power transformer outage is often counted in millions of pounds for output loss only, not to say of the costs associated with equipment repair or replacement. Because of this economic incentive, preventive tests and on-line monitoring are of benefit to predict incipient fault conditions, and to schedule outage, maintenance and retirement of transformers.

Fault conditions in a power transformer can be detected in several ways. One method is based on detection of PDs which are generated due to fault conditions related to moisture, cavities in the solid insulation, metallic particles, and gas bubbles. A significant increase either in the PD level or in the developing rate of PD activity can provide an early indication of an incipient fault condition.

PD monitoring methods generally fall into two categories: the electrical method and the acoustic method. According to the sensors they are using, both of them can be further classified into two sub-categories: non-intrusive monitoring and intrusive monitoring. Non-intrusive monitoring uses sensors installed outside the transformers, such as coupling capacitors, high frequency current transformers, piezoelectric type accelerometers, etc. Intrusive monitoring must put the PD sensors inside the transformer and the implementation is thus more dangerous and expensive.

Although experts all over the world have investigated methods for PDs detection and classification for more than 50 years, until now there is still not available a universal method for this purpose.

In view of the foregoing, the work presented in this thesis has been focused to the investigation and development of a non-invasive approach for detection and interpretation of the PD dedicated to power transformers. The major focus of the work presented herein has been on developing the non-invasive method, based on broadband measurements. The work can be divided into three main parts: the design of broadband measuring set-up, taking measurements and the analysis and interpretation of captured data. Since remote radiometric measurements have not been investigated in any great detail in the past, fingerprints of different discharges in power transformers obtained in such a manner are not available. Thus, as the first stage, the formation of groups obtained from captured PD data was crucial and this included the implementation of a number of methods for grouping captured data. Moreover, the separation of PDs originating from a transformer and discharges originating from the surrounding is achieved.

This chapter briefly reviews all the major studies that have been performed in this research project and a number of significant achievements that have been attained. Finally, possible future work with regard to this research is also proposed at the end of this chapter.

14.2. Review of Major Studies and Achievements Attained in This Research Project

Throughout the three years of research on this project, various major studies have been performed, as reported in this thesis. More importantly, several important achievements have been attained as a result. The following sections will provide a brief review of these major studies and the achievements that have been attained.

14.2.1. A design of non-invasive PD measurements dedicating to power transformers

The use of remote radiometric measurements for identifying PDs occurring in power transformers is implemented in this research. The main advantage of this technique is that it does not require a galvanic connection to the transformer. Therefore the measurements can be performed during transformer operation. The first trial measurements were based on narrow-band measurements, consisting of different types of antennas, receiver, oscilloscope and laptop. Three types of antennas were used at different locations, but antennas were not placed around the transformer. The findings of the first trials clearly showed that wide band measurements should be used for this purpose rather than narrow band. Phase resolved measurements are not adequate and therefore, it is necessary to record the PD signals and analyze their shapes.

The second trial measurements were based on wideband measurements. Antennas were placed around the transformer. Antennas were directly connected to the very advanced oscilloscope and the laptop. Analysis of results for the second trial showed that the set-up was appropriate and useful for measuring PDs in transformers.

14.2.2. Dimensionality reduction of captured data

Thousands of signals were captured in the substation. Additionally, each signal was represented by 5000 samples. For this reason, it was not possible to manually analyse captured data. In view of the fact that the computer memory as well as Matlab program used for the analysis of captured data have limitations, this part of the research was focused on finding the appropriate method for dimensionality reduction.

Amongst the many feature extraction methods, PCA and method of moments were chosen as the methods for dimensionality reduction of PD signals. The aim of the PCA is to find a set of M orthogonal vectors in data space that account for as much as possible of the data's variance. Projecting the data from its original N-dimensional space onto the M-dimensional subspace spanned by these vectors then performs a dimensionality reduction that often retains most of the intrinsic information in the data.

The method of moments enables representation of captured signals by only two parameters, i.e. "equivalent time-length" and "equivalent bandwidth". The strength of the method lies in its visualization in a 2-dimensional space.

14.2.3. Application of SOM on PD data

A study of DM techniques is necessary to group different signals captured in the substation. In this work, a DM technique known as the SOM, has been applied for the analysis and interpretation of captured PD data. Before applying the SOM, the dimensionality reduction based on PCA was performed. It has been illustrated herein that the SOM is capable of grouping different PD pulses and representing these revealed features in a discernible way. The proposed approach has the advantage of allowing the "visualisation" of obtained groups in the u-matrix. The disadvantage of this method is that the number of input variables has a big impact on the formation of groups. Therefore sufficient number of input variables must be taken. Additionally, the appropriate SOM must be found for different number of input variables, and this can be time consuming.

14.2.4. Application of CA on PD data

Another method for grouping data into clusters was applied in this work. Among the CA techniques, the hierarchical methods were chosen since they give greater insight into the overall structure. Investigations based on Euclidian and Mahalanobis distance measures and Ward and Average linkage algorithms were thus performed on PD data preprocessed by PCA. In this respect, Ward linkage algorithm proved to be the best method for clustering PD data. It was found that the visualization of CA method was possible only if three input variables were selected for CA. Through this approach, data can be presented in a 3-dimensional space. CA

was applied to different number of input variables and it was found that the formation of groups was not very much influenced by the number of input variables. It would thus appear that CA has advantages over SOM for PD data and the latter was therefore adopted in this work.

14.2.5. Analysis of signals originating from different defect models in the HV Laboratory

In order to show that different type of PDs emitted EM signals with different frequency spectra, some sources of PDs were generated in the HV Laboratory. As expected, based on the frequency spectra of the signals, the characteristics of discharges can be identified and it thus becomes possible to differentiate discharges emanating from oil and discharges emanating from air. Furthermore, tests performed in the Laboratory confirmed that some signals captured at the Northfleet substation, originated from the power transformer itself.

14.2.6. Method of moments applied to the PD data

Another feature extraction method implemented on PD measurements was the method of moments. The advantage of this method lies in the transformation of a PD signal into time-frequency plane. Analysis of the results showed, that transformation of PDs into time-frequency plane is not appropriate for PD measurements acquired during this research, since signals of different PD pulses mapped on to the same σ_T and σ_F values. Therefore, σ_T and σ_F values are not sufficient to discriminate phenomena due to different PD sources.

14.2.7. Final conclusions

Since remote radiometric measurements for detecting PD behaviour in power transformer had not been performed in the past, it was not possible to know what kind of signals were to be expected when the research was first started. When measurements were performed at the substation, there was no clear picture as to how many different signals could be expected. Another problem was how to differentiate between signals originating from the transformers from those originating from transformer surroundings.

The origin of captured signals in the substation was another unknown entity. However, this made it possible to analyse a bunch of signals by clustering/groupings. The task was to divide captured signals into several coherent groups. The information regarding the PD shape and its frequency spectrum was also not known. Thus it was not possible to know how many groups of signals could be expected. When analysing signals, the main focus was to separate signals originating from the oil, i.e. PDs emanating from the transformer, from those originating from transformer surroundings. Tests in the HV Laboratory proved that it was possible to separate PD signals emanating from oil and signals from air, based on their frequency content.

Grouping of signals was based on two different DM methods, the SOM and the CA, respectively. The presented results show that CA gives better performance then SOM. Probably the SOM would have given better results if the number of groups were known and the task was simply confined to the classification problem. In the absence of any knowledge of the actual number of groups, it was thus more appropriate to use the CA method for the presented case. Before applying the CA it was important to choose the appropriate linkage algorithm. For clustering PD signals it has been shown that the Ward's linkage algorithm is the best suited, since it enables distinct groups to be formed with a relatively high number of signals with similarities within each group rather than seek individuals.

In order to construct a reliable and efficient system for PD measurement, dedicated to power transformers a considerable effort needs to be expended. For example, groups of PD signals emanating from the transformer must be known, i.e. the shape of pulses and their frequency spectra. In essence, every transformer must be treated on an individual basis and this would involve performing tests on transformers with different ratings and different voltage levels. Setting up a comprehensive data base for different transformers is of crucial importance.

14.3. FUTURE WORK

The work presented in this thesis is the beginning of a big challenge to create a system for PD detection and identification based on remote radiometric measurements in a real life situation. The presented results show that non-invasive measurements are potentially quite attractive and hence the suggestions for future work are:

- First of all extensive tests in the Laboratory should be performed. It is important to create all conceivable defects that can cause PDs in transformers, this will enable "clean" fingerprints of different defects to be ascertained.
- It is vitally important to achieve repeatability of measurements in a Laboratory, since this is the only way to obtain the exact fingerprints. If possible, it is desirable to create two different PDs at the same time to see the influence of coincident pulses on the fingerprints.
- The next step is to investigate the influence of different sources of noise, since only this is the only way of classification of different types of noise. Additionally, the surface discharges of the bushing should be studied.
- In the following step techniques for noise elimination should be studied and applied to the captured signals.

- When good knowledge about different discharges is obtained, the on-site PD measurement results should be performed. Different transformers, i.e., step-up transformers in a power plant, transformers installed in transmission system as well as distribution transformers should be chosen. Moreover, transformers containing OLTC inside transformer tank and transformers with separate OLTC should be tested. The influence of transformer design should be investigated too.
- The location of a PD source is also very valuable information; therefore determination of PD sources based on triangulation should be performed.
- Perhaps a special designed antenna exclusively for PD capture should be developed.
- Setting up of good and credible databases of different discharge fingerprints occurring in power transformers of different voltage levels should be put in place.

15. REFERENCES

[1] K. Karsai, D. Kerenyi, L. Kiss: "Large Power Transformers", Akademiai Kiado, Budapest 1987.

[2] L. Petterson, N.L. Fantana, U. Sundermann: "Life assessment: Ranking of power transformers using condition based evaluation. A new approach", CIGRE 1998, pp. 122-204.

[3] WG 12.09: "Lifetime evaluation of transformers", Electra, No.150, October 1993.

[4] T. Hoshino, K. Kato, N. Hayakawa, and H. Okubo, "A novel technique for detecting electromagnetic wave caused by partial discharge in GIS," IEEE Transactions on Power Delivery, vol. 16, pp. 545-51, 2001.

[5] Y. C. Leung and J. M. K. MacAlpine, "Initial experience with the partial discharge monitoring of high-voltage motors," Electric Power Systems Research, vol. 61, pp. 33-40, 2002.

[6] C. Hudon, A. Contin, M. Bélec, A. Cavallini, D.N. Nguyen, G.C. Montanari and M. Conti: "Evolution in automatic phase resolved partial discharge pattern recognition for rotating machine diagnosis," XIIIth International Symposium on High Voltage Engineering, Netherlands 2003.

[7] B.T. Phung, Z. Liu, T.R. Blackburn and R.E. James , "Recent Developments for On-Line Partial Discharge Detection in Cables", 2001 Australasian Universities Power Engineering Conference, Curtin University of Technology, 413-418, ISBN 176067068X, Curtin University of Technology, Perth, Australia.

[8] T. Bengtsson, H. Kols, B. Jönsson: "Transformer PD diagnosis using acoustic emission technique", ISH 97, Montreal, Canada.

[9] L.E. Lundgaard: "Partial discharge XIII: Acoustic partial discharge detection-fundamental considerations", IEEE Electrical Insulation Magazine, vol.8, no.4, July-Aug. 1992, pp.25-31.

[10] L.E. Lundgaard: "Partial discharge XIV: Acoustic partial discharge detection-practical application", IEEE Electrical Insulation Magazine, vol.8, no.5, Sept.-Oct. 1992, pp.34-43.

[11] P. M. Eleftherion: "Partial discharge XXI: Acoustic emission-based PD source location in transformers", IEEE Electrical Insulation Magazine, Vol. 11, No. 6, November/December 1995.

[12] IEC 60270: High-voltage test techniques - Partial Discharge Measurements, 2000.

[13] A.J.M. Pemen, W.R Rutgers, T.J.M van Rijn, Y.H. Fu: "On-line partial discharge monitoring of HV components", Eleventh International Symposium on High Voltage Engineering, London, UK (Conf. Publ. No. 467), IEE. Part vol.5, 1999, pp. 136-139.

[14] R. Schwarz, M. Muhr, S. Pack: "Partial discharge detection and localization in transformers" XIIIth International Symposium on High Voltage Engineering, Netherlands 2003.

[15] A. M. Emsley, X. Xiao, R. J. Heywood, and M. Ali, "Degradation of cellulosic insulation in power transformers. Part 3: effects of oxygen and water on ageing in oil", IEE Proceedings-Science, Measurement and Technology, Vol. 147, No. 3, pp. 115-19, 2000.

[16] M. A. Salam, H. Anis, A.E. Morshedy, R. Radwan, High Voltage Engineering, Theory and Practice, 2nd Edition, Marcel Dekker, Inc., 2000.

[17] L. Centurioni and G. Coletti, "Transformer Insulation", Wileys Encyclopedia of Electrical and Electronics Engineering online, 2000.

[18] T. O. Rouse, "Mineral insulating oil in transformers", IEEE Electrical Insulation Magazine, Vol. 14, No. 3, pp. 6-16, 1998.

[19] S. M. Gubanski, P. Boss, G. Csepes, V. D. Houhanessian, J. Filippini, P. Guuinic, U. Gafvert, V. Karius, J. Lapworth, G. Urbani, P. Werelius, and W. S. Zaengl, "Dielectric response methods for diagnostics of power transformers", Electra, No. 202, pp. 23-34, June, 2002.

[20] I. Kobal, S. Vizintin, M. Jerele: "Kontrola izolacije energetskih transformatorjev", Power Institute Milan Vidmar, Ljubljana, 1987 (In Slovenian language).

[21] C. Sepes, I. Hamos, I. Kispal, J. Schmidt, A. Bognar: " A DC Expert system (RVM) for checking the refurbishment efficiency of high voltage oil-paper insulating system using polarization spectrum analysis in range of long-time constants", CIGRE 1994, 12-206.

[22] E. Kueffeland, M. Abdullah: "High-voltage engineering", Pergamon Press, 1970, pp.313-317.

[23] S. Birlasekaran, F. Fetherston: "Off/on-line FRA condition monitoring technique for power transformers", IEEE Power Engineering Review, August 1999, pp.54-56.

[24] IEC Publication 60422: Supervision and maintenance guide for mineral insulating oils in electrical equipment, 1989.

[25] IEC Publication 60559: Interpretation of analysis of gases in transformers and other oil-filled electrical equipment in service, 1999.

[26] A. Möllmann, B. Pahlavanpour: "New guidelines for interpretation of dissolved gas analysis in oil-filled transformers, Electra, No. 186, October 1999.

[27] M. Ali, C. Eley, A.M. Emsley, R. Heywood, X. Xaio: "Measuring and understanding the ageing of kraft insulating paper in power transformers", IEEE Electrical Insulation Magazine, vol.12, no.3, May-June 1996, pp.28-34.

[28] H.P. Gasser, J. Huser, C. Krause, V. Dahinden, A.M. Emsley: "Determining the ageing parameters of cellulosic insulation in a transformer", Eleventh International Symposium on High Voltage Engineering (Conf. Publ. No.467). IEE. Part vol.4, 1999, pp.143-7.

[29] R.M. Morais, W.A. Mannheimer, M. Carballeira, J.C. Noualhaguet: "Furfural analysis for assessing degradation of thermally upgraded papers in transformer insulation", IEEE Transactions on Dielectrics & Electrical Insulation, vol.6, no.2, April 1999, pp.159-63.

[30] J.P. Steiner: "Partial discharge – Part IV: Commercial PD testing", IEEE Electrical Insulation Magazine, Jan/Feb. Vol.7, No.1, 1991, pp.20-33.

[31] http://www.uesystems.com/products/up2000.htm

[32] C. Bengtsson: "Status and Trends in Transformer Monitoring", IEEE Trans. on Power Delivery, Vol. 11, No. 3, July 1996, pp. 1379-1384.

[33] Special report "On-line Transformer Monitoring", Electrical World, October 1995, Vol. 209, No. 10, pp. 19-26.

[34] J. Douglas: "The maintenance revolution", EPRI Journal, May/June 1995, pp. 8-15.

[35] M. Eby, R. Bush: "Maintenance management techniques for the future", Transmission and Distribution World, August 1996, pp. 94-103.

[36] T. Leibfried: "Online monitors keep transformers in service", IEEE Computer applications in Power, July 1998, pp. 36-42.

[37] CIGRE WG 05: "An international survey on failures in large power transformers in service", Electra, No.88, 1983.

[38] GE: http://www.gepower.com

[39] Morgan Schafer: http://www.morganschaffer.com/

[40] Serveron: http://www.micromonitors.com/

[41] Power Diagnostix Systems GmbH: http://www.pd-systems.com/

[42] P.M. Eleftherion: "Partial Discharge XXI: Acoustic emission-based PD source location in transformers", IEEE Electrical Insulation Magazine, November/December 1995-Vol. 11, No.6.

[43] Physical Acoustic Corporation: http://www.pacndt.com/featbasd.html

[44] Cuttler-Hammer: http://www.partial-discharge.com/

[45] Weschler instruments: http://www.winst.com/Advantage/brochure.html

[46] Luxtron Corporation: http://www.luxtron.com/

[47] Doble Engineering: http://www.doble.com/

[48] Panametrics: http://www.panametrics.com/

[49] GE Harley: http://www.gepower.com/reuterstokes/harley/monitoring.html

[50] Barrington Consultants Incorporated: http://www.barringtoninc.com/

[51] I. Fofana, V. Wasserberg, H. Borsil, and E. Gockenbach, "Retrofilling conditions of high voltage transformers", IEEE Electrical Insulation Magazine, Vol. 17, 2001.

[52] S. Teenbohlen, F. Figel: "On-line condition monitoring on power transformers", IEEE summer meeting 2000, 0-7803-5938.

[53] ABB Transformers: http://www.abb.se/setfo/

[54] Siemens: http://www.transformer-service.com/en/products_monitoring_main.htm

[55] Gulski E.: "Digital analysis of partial discharges", IEEE Transactions on Dielectrics & Electrical Insulation, vol.2, no.5, Oct. 1995, pp.822-37.

[56] F. H. Kreuger: "Partial discharge detection in high-voltage equipment", Butterworths, 1989, ISBN 0-408-02063-6.

[57] P. Pakonen, A. Lehtio, K. Kannus, E. Lakervi: "RFI measurements as a diagnostic technique for medium voltage transformers", International Conference on Partial Discharge, London, UK (Conf. Publ. No. 378), IEE. 1993, pp. 115-116.

[58] Task Force 15.01.04: "Partial discharges in Transformer Insulation", CIGRE 2000, 15-302.

[59] C. Bengtsson, L. Pettersson, U. Gäfvert, T. Bengtsson, G. Frimpog, J. Fuhr, N. L. Fantala: "Modern field diagnostic techniques for power transformers", IEEMA Trafotech-98, Mumbai, India.

[60] D. Chu, L.J. Savio, S.R. Lindgren, H.R. Moore, S.L. Nilsson: "Application of in-service monitoring system during tests at consolidated Edison's Ramapo station", CIGRE 1998, 12-205.

[61] T.R. Blackburn, B.T. Phung, R.E. James: "Optical fibre sensor for partial discharge detection and location in high-voltage power transformer" Sixth International Conference on Dielectric Materials, Measurements and Applications (Conf. Publ. No.363). IEE. 1992, pp.33-6.

[62] J. Deng, H. Xiao, W. Huo, M. Luo, R. May, A. Wang, Y. Liu: "Optical fiber sensor-based detection of partial discharges in power transformers", Optics & Laser Technology, Volume 33, Issue 5, July 2001, Pages 305-311.

[63] WG 21.03 (F.H. Kreuger): "Recognition of discharges", Electra, No1 11, 1970, pp.61-79.

[64] T. Eriksson, M.Leijon, C. Bengtsson: "PD on-line monitoring of power transformers", Stockholm Power Tech International Symposium on Electric Power Engineering. IEEE Part vol.6, 1995, pp. 101.1-101.5.

[65] S. Tenbohlen, D. Uhde, J.Poittevin, U. Sundemann, H. Borsi, P. Werle, H. Matthes: "Enhanced diagnosis of power transformers using on- and off-line methods: results, examples and future trends", CIGRE 2000, 12-204.

[66] J.P. Bolhuis, E. Gulski, J.J. smit, T. Grun, M. Turner: "The comparison of conventional and VHF PD detection methods for power transformers", Eleventh International Symposium on High Voltage Engineering, London, UK (Conf. Publ. No. 467), IEE. Part vol.5, 1999, pp. 49-52.

[67] E. Gulski, F.H. Kreuger.: "Computer-aided recognition of discharge sources", IEEE Transactions on Electrical Insulation, vol.27, no.1, Feb. 1992, pp.82-92.

[68] F.H. Kreuger, E. Gulski, A. Krivda: "Classification of partial discharges", IEEE Transactions on Electrical Insulation, vol.28, no.6, Dec. 1993, pp.917-31.

[69] E. Gulski: "Computer-aided measurement of partial discharges in HV equipment", IEEE Transactions on Electrical Insulation, vol.28, no.6, Dec. 1993, pp.969-83.

[70] E. Gulski: "Digital analysis of partial discharges", IEEE Transactions on Electrical Insulation, vol.2, no.5, Oct. 1995, pp.822-837.

[71] C. Chang, Q. Su: "Analysis of PD patterns with integrated distribution parameters using modular neural networks" PowerCon 2000. 2000 International Conference on Power System Technology. Proceedings (Cat. No.00EX409). IEEE. Part vol.1, 2000, pp.259-64 vol.1.

[72] R.E. James, B.T. Phung: "Development of computer-based measurements and their application to PD pattern analysis", IEEE Transactions on Dielectrics & Electrical Insulation, vol.2, no.5, Oct. 1995, pp.838-56.

[73] E. Gulski: "Discharge pattern recognition in high voltage equipment", IEE Proc.-Sci.Meas. Technol., Vol. 142, No.1, January 1995, pp.51-61.

[74] H.G. Kranz, R. Krump: "Partial discharge diagnosis using statistical optimisation on a PC-based system", IEEE Transactions on Electrical Insulation, vol.27, no.1, Feb. 1992, pp.93-98.

[75] L. Satish, B. Gurugaj: "Use of hidden Markov models for partial discharge pattern classification", IEEE Transactions on Electrical Insulation, Vol.28, no.2, April 1993, pp.172-182.

[76] M.M.A. Salama, R. Bartnikas: "Fuzzy logic applied to PD pattern classification", IEEE Transactions on Electrical Insulation, Vol.7, no.1, Feb. 2000, pp.118-123.

[77] N. Huzomi, T. Okamoto, T. Imajo: "Discrimination of partial discharge patterns using a neural network", IEEE Transactions on Electrical Insulation, Vo.27, no. 3, June1992, pp.550-556.

[78] A. Gross, S. Happe, H.G. Kranz: "Time-resolved signal recognition with a neural network hardware using a decentralized PC-system" Fourth European Congress on Intelligent Techniques and Soft Computing Proceedings, EU FIT '96. Verlag Mainz. Part vol.3, 1996, pp.1561-4 vol.3.

[79] T. Eriksson, M.Leijon, C. Bengtsson: "PD on-line monitoring of power transformers", Stockholm Power Tech International Symposium on Electric Power Engineering. IEEE Part vol.6, 1995, pp. 101.1-101.5.

[80] H.G. Kranz, A.Lapp, D. Aschenbrenner: "Automated partial discharge diagnosis system WinTED 2.0", http://www.lfh.uni-wuppertal.de

[81] P.H. F. Morshuis: "Time-resolved discharge measurements", International Conference on Partial Discharge (IEE Conf. Publ. No. 378), 1993, pp.43-46.

[82] C.H. Peck, P.J. Moore: "Source Identification and Analysis of Wideband Electromagnetic Interference Generated by Switch-gear in a High Voltage Substation", EMC2000, Paper 2A7, 5p, York.

[83] W.R. Rutgers, P. van den Aardweg, A. Lapp, H.G. Kranz: "Transformer PD measurements: Field experience and automated defect identification", XIII Int. Conf. on gas discharges and their applications, Glasgow, 2000.

[84] W. R. Rutgers, Y. H. Fu: "UHF PD-Detection in a Power Transformer", ISH97, Montreal, Canada, 1997.

[85] M. Kantardzic, Data mining; Concepts, models, methods and algorithms, A John Wiley & Sons, Inc., Publication, 2003.

[86] C. Olaru, L. Wehenkel: "Data mining", IEEE Computer Application in Power, July 1999, pp.19-25.

[87] R. A. Johnson, E. W. Wichern: "Applied Multivariate Statistical Analysys", Fifth Edition, Prentice Hall, 2002.

[88] A. Hyvärinen, J. Karhunen, E. Oja: "Independent Component Analysis", John Wiley & Sons,Inc, 2001.

[89] D. Lowe and A.R. Webb: "Optimized Feature Extraction and the Bayes Decision in Feed-Forward Classifier Networks", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 13, no. 4, pp. 355-264, Apr. 1991.

[90] A. Ultsch: "Knowledge Extraction from Self-Organizing Neural Networks", Information and Classification, 1993, pp. 301-306.

[91] L.E. Franks, Signal Theory, Prentice-Hall, Inc., 1969.

[92] T. Kohonen: "Self-Organizing Maps (2nd Edition)" (Springer-Verlag, 1997).

[93] J. Vesanto: "Data Mining Techniques Based on the Self-Organizing Map", Thesis for the degree of Master of Science in Engineering, Department of Engineering Physics and Mathematics, Helsinki University of Technology, 26th May 1997.

[94] A. Ultsch: "Self-Organizing Neural Networks for Visualisation and Classification", Information and Classification, 1993, pp. 307-313.

[95] A. Ultsch: "Knowledge Acquisition with Self-Organizing Neural Networks", Proceedings of the 1992 International Conference, 1992, Vol. 1, pp. 735-739.

[96] Neural Network Research Centre, Helsinki University of Technology: "On-line Documentation for SOM Toolbox V2.0 Beta for Matlab 5.0", (http://www.cis.hut.fi/projects/somtoolbox/package/docs2/somtoolbox.html).

[97] SOM_PAK (http://www.cis.hut.fi/research/som_lvq_pak.shtml).

[98] SOM_TOOLBOX (http://www.cis.hut.fi/projects/somtoolbox/).

[99] L. Kaufman, P.J. Rousseeuw: "Finding groups in data. An introduction to cluster analysis", John Wiley & Sons, 1990.

[100] L. Lebart, A. Morineau: "Multivariate descriptive statistical analysis", John Wiley & Sons, 1984.

[101] W.R. Dillon, M. Goldstein: "Multivariate analysis; Methods and applications", John Wiley & Sons, 1984.

[102] P.J. Rousseeuw: "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis", Journal of Computational and Applied Mathematics, 20(1987), pp. 53-65.

[103] E.M. Lalitha, L.Satish: "Wavelet analysis for classification of multi-source PD patterns", IEEE Transactions on Dielectrics and Electrical Insulation, Vol.7, no.1, Feb. 2000, pp.40-47.

[104] A. Cavallini, A. Contin, G. C. Montanari, F. Puletti, "Advanced PD inference in on-field measurements. Part 1: noise rejection", IEEE Trans. on Dielectrics and Electrical Insulation, Vol. 10, n. 2, April 2003, pp. 216-224.

[105] A. Contin, A. Cavallini, G. C. Montanari, G. Pasini, F. Puletti, "Digital Detection and Fuzzy Classification of Partial Discharge Signals", IEEE Trans. on DEI, Vol. 9, no. 3, pp.335-348, June 2002.

[106] A. Cavallini, A. Contin, G.C: Montanari, F. Puletti, "A new approach to diagnosis of solid insulation systems based on PD signal inference", IEEE Electrical Insulation Magazine, Vol. 19, no. 32, pp. 23-30, April 2003.

[107] E. Howells, E.T. Norton. "Detection of Partial Discharges in Transformers Using Acoustic Emission Techniques". IEEE Trans.PAS-97, No. 5, Sept/Oct 1978. p. 1538-1549.

16. APPENDIX

16.1. RESULTS OF THE MEASUREMENTS OBTAINED BY YAGI 2

Figures 16.1 to 16.4 present four main groups of signals and their frequency spectra measured by Yagi 2 pointing at phase A on 275 kV side of the transformer (left) and signals detected by the KEMA probe (right).

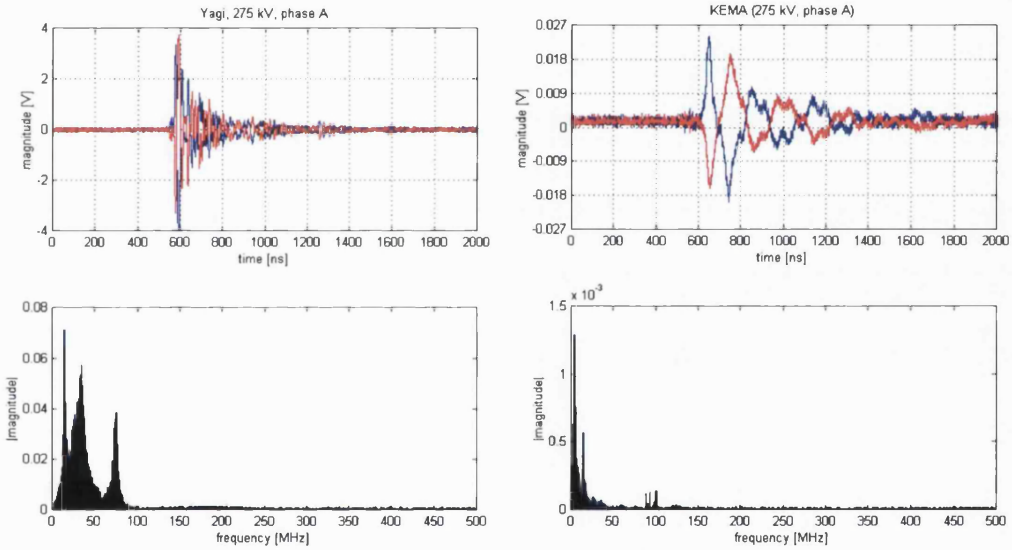


Figure 16.1: 1st group of signals measured by Yagi antenna in phase A on 275 kV side of the transformer and the KEMA probe

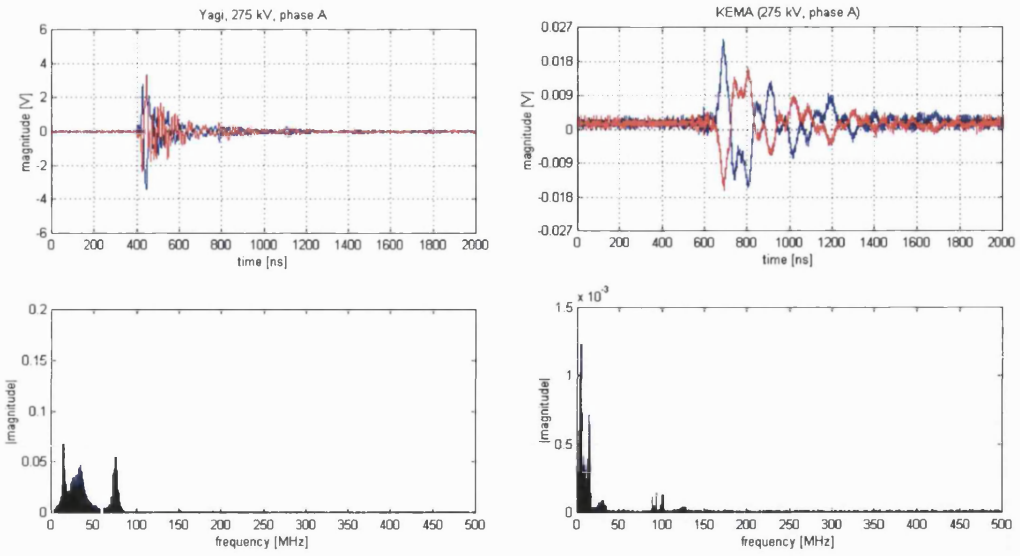


Figure 16.2: 2nd group of signals measured by Yagi antenna in phase A on 275 kV side of the transformer and the KEMA probe

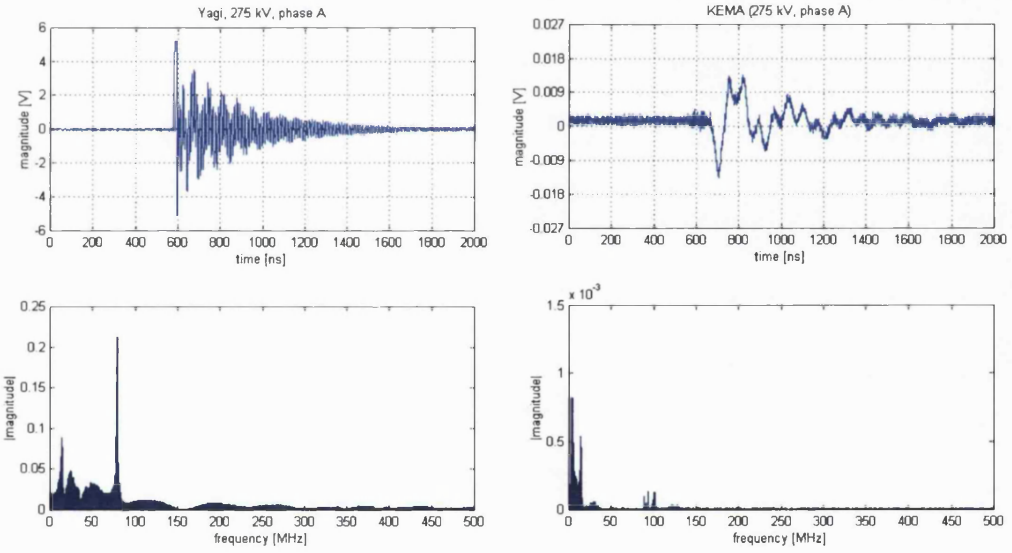


Figure 16.3: 3rd group of signals measured by Yagi antenna in phase A on 275 kV side of the transformer and the KEMA probe

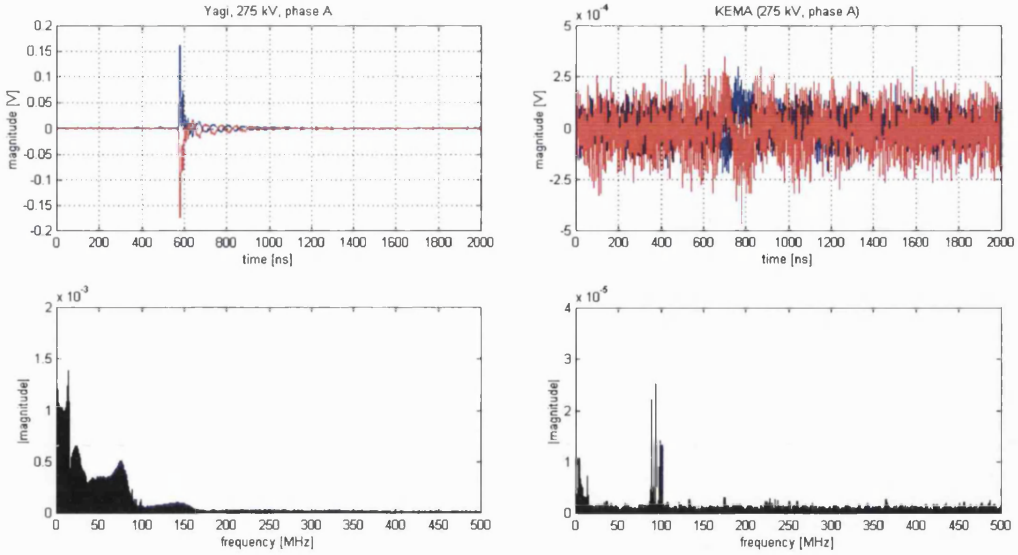


Figure 16.4: 4th group of signals measured by Yagi antenna in phase A on 275 kV side of the transformer and KEMA probe

Figures 16.5 to 16.7 contain the number of pulses with regard to the voltage phase angle. Figure 16.5 illustrates distribution of all groups of discharges together. Figure 16.6 presents distribution of the number of all discharges with the green colour and with the black colour discharges of the 1st and 2nd group. Figure 16.7 shows the situation for the 3rd (left side) and 4th (right side) group of discharges. Discharges of first two groups occur between 350 and 50° and between 165 and 220° while discharges of the 3rd group take place between 165-205° while the 4th group between 0-35° and 180-220°.

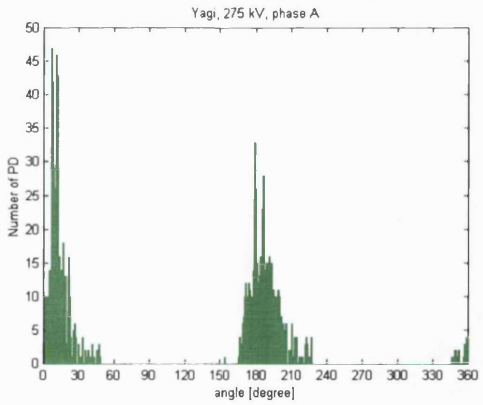


Figure 16.5: Distribution of the number of pulses regarding the voltage phase angle

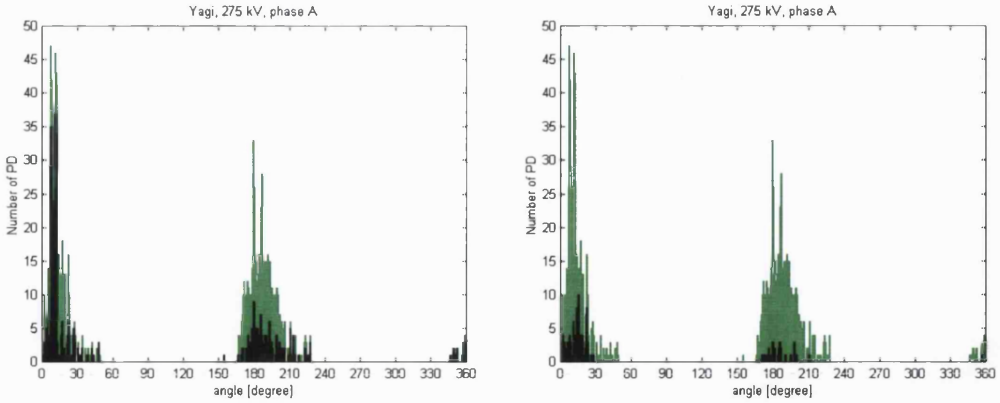


Figure 16.6: Distribution of the number of pulses regarding the voltage phase angle for the 1st (left graph) and the 2nd (right graph) group of signals

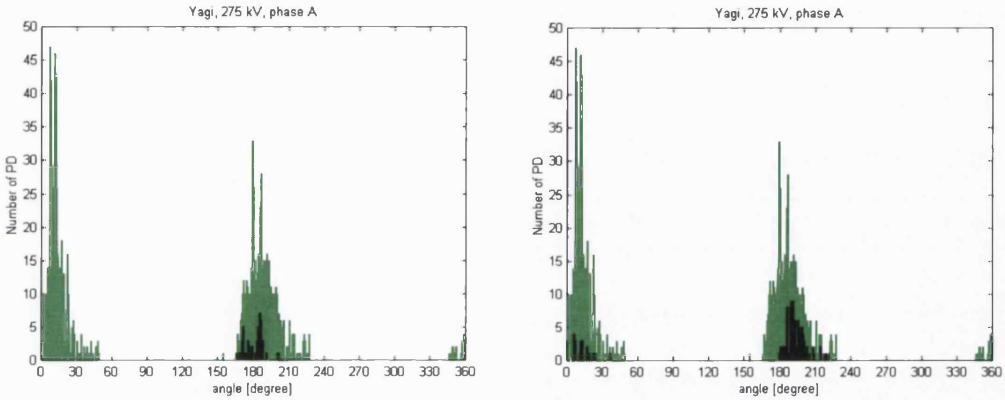


Figure 16.7: Distribution of the number of pulses regarding the voltage phase angle for the 3rd (left graph) and the 4th (right graph) group of signals

16.2. RESULTS OF THE MEASUREMENTS OBTAINED BY YAGI 3

Figures 16.8 and 16.9 show signals and their frequency spectra measured by Yagi 3 pointing at phase C on 400 kV side of the transformer (left) and signals detected by the KEMA probe (right).

Figures 16.10 and 16.11 show the number of pulses with regard to the voltage phase angle. Figure 16.10 illustrates distribution of all groups of discharges together. Figure 16.11 presents distribution of the number of all discharges with the green colour and with the black colour discharges of the 1st (left graph) and 2nd group (right graph). Discharges of the second group occurred only in the positive halve of the reference voltage.

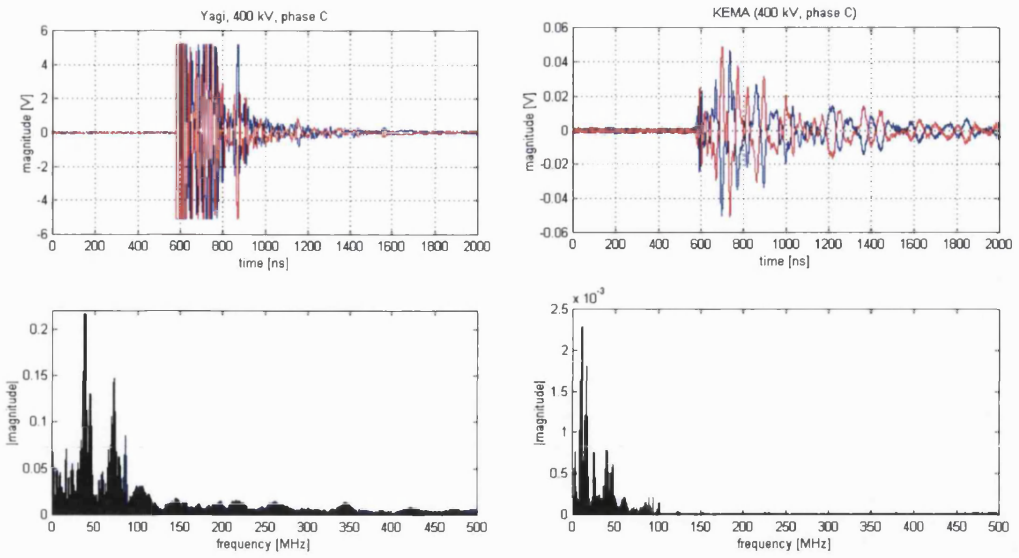


Figure 16.8: 1st group of signals measured by Yagi antenna in phase C on 400 kV side of the transformer and the KEMA probe

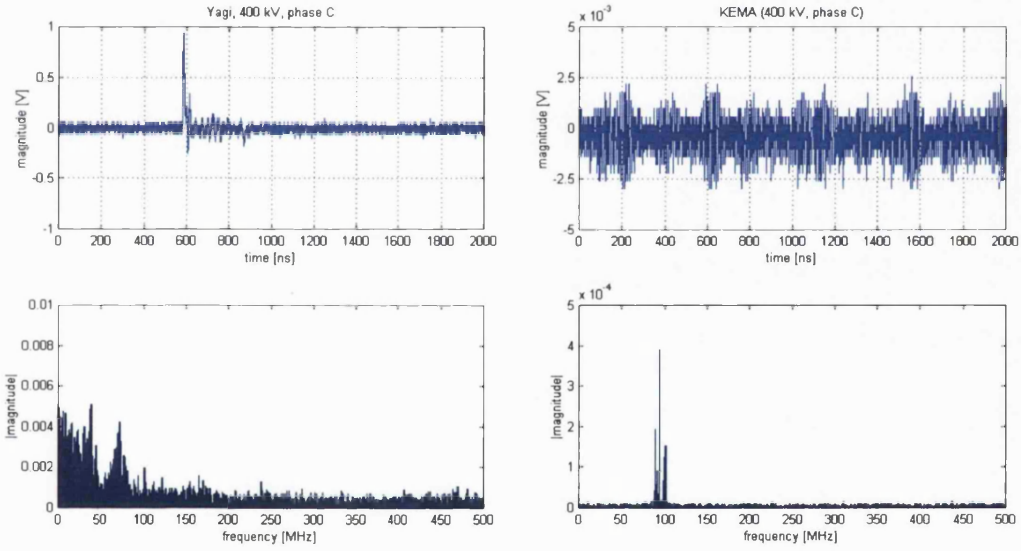


Figure 16.9: 2nd group of signals measured by Yagi antenna in phase C on 400 kV side of the transformer and the KEMA probe

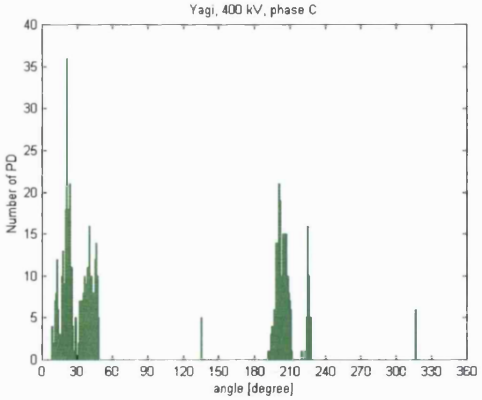


Figure 16.10: Distribution of the number of pulses regarding the voltage phase angle

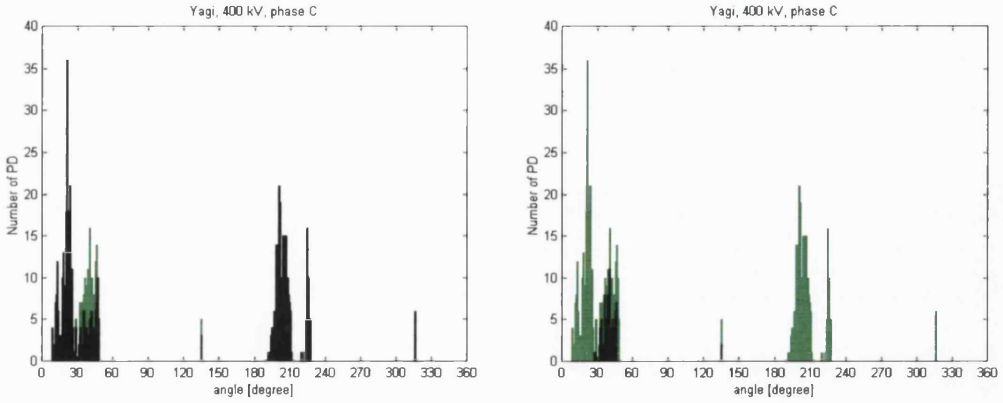


Figure 16.11: Distribution of the number of pulses regarding the voltage phase angle for the 1st and the 2nd group of signals

16.3. RESULTS OF THE MEASUREMENTS OBTAINED BY H2

Figures 16.12 and 16.13 show two main groups of signals and their frequency spectra measured by the helical antenna H2, placed close to phase A on 275 kV side of the transformer. For the same shape of signal obtained by the helical antenna, the KEMA probe detected two main types of signals as shown in Figure 16.12. The shapes of both groups of signals are quite similar; the 2nd group of signals detected by helical antenna are smaller in magnitude. The signals were triggered by helical antenna signals.

Figure 16.14 shows the number of pulses with regard to the voltage phase angle. The distribution of discharges is between 0 - 50° and 175 - 230°.

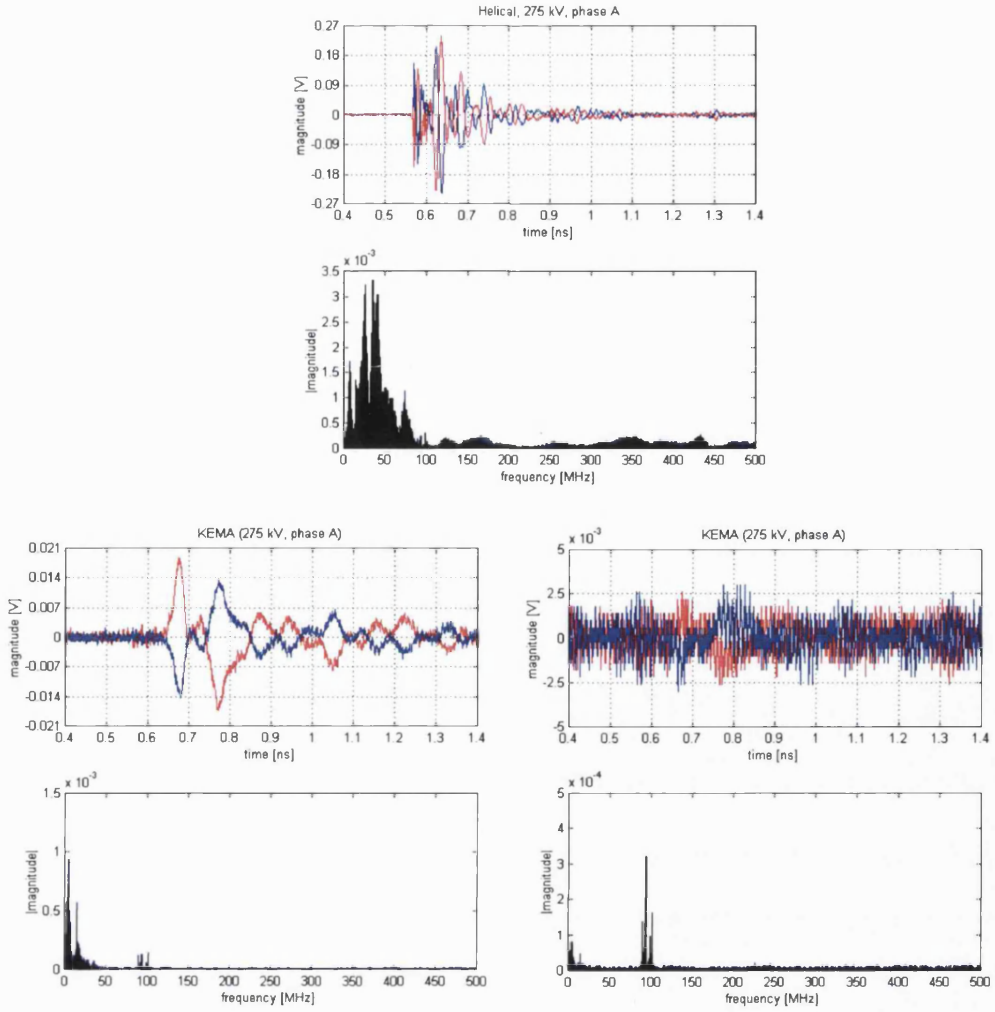


Figure 16.12: 1st group of signals measured by the helical antenna H2 and two types of the KEMA probe signals

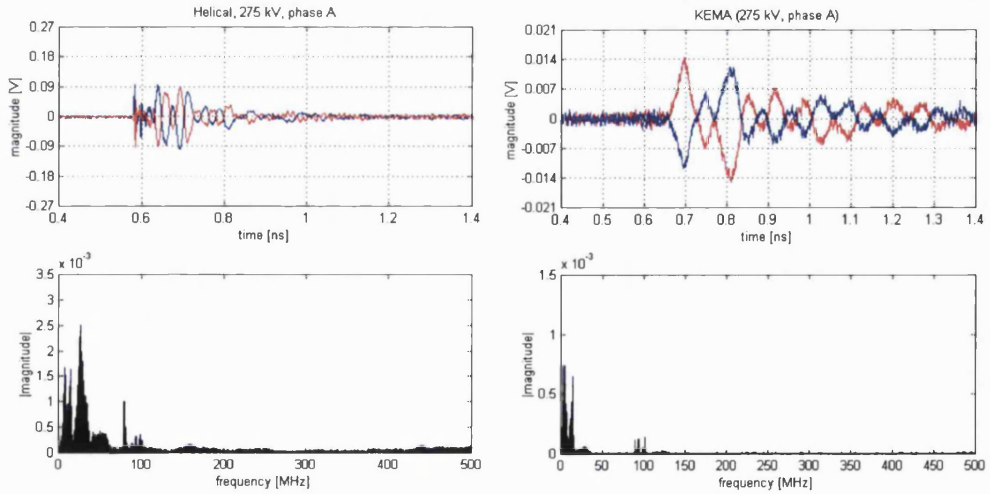


Figure 16.13: 2nd group of signals measured by the helical antenna H2 and the KEMA probe

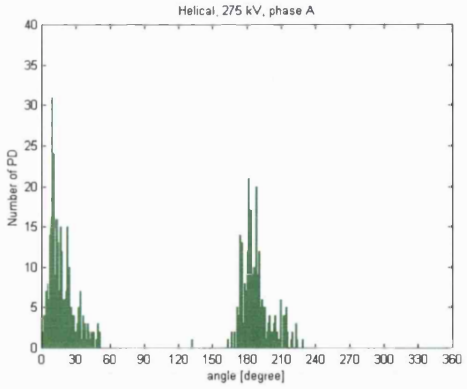


Figure 16.14: Distribution of the number of pulses regarding the voltage phase angle for H2

16.4. RESULTS OF THE MEASUREMENTS OBTAINED BY H3

Signals captured by the helical antenna H3, placed closed to phase A on 400 kV side only contain signals of one type while signals obtained by the KEMA probe can be divided into three groups as shown in Figure 16.15.

Figure 16.16 presents the number of pulses with regard to the voltage phase angle. The distribution of the signals is between 15 - 50° and 190 - 230° and few discharges occur at 315°.

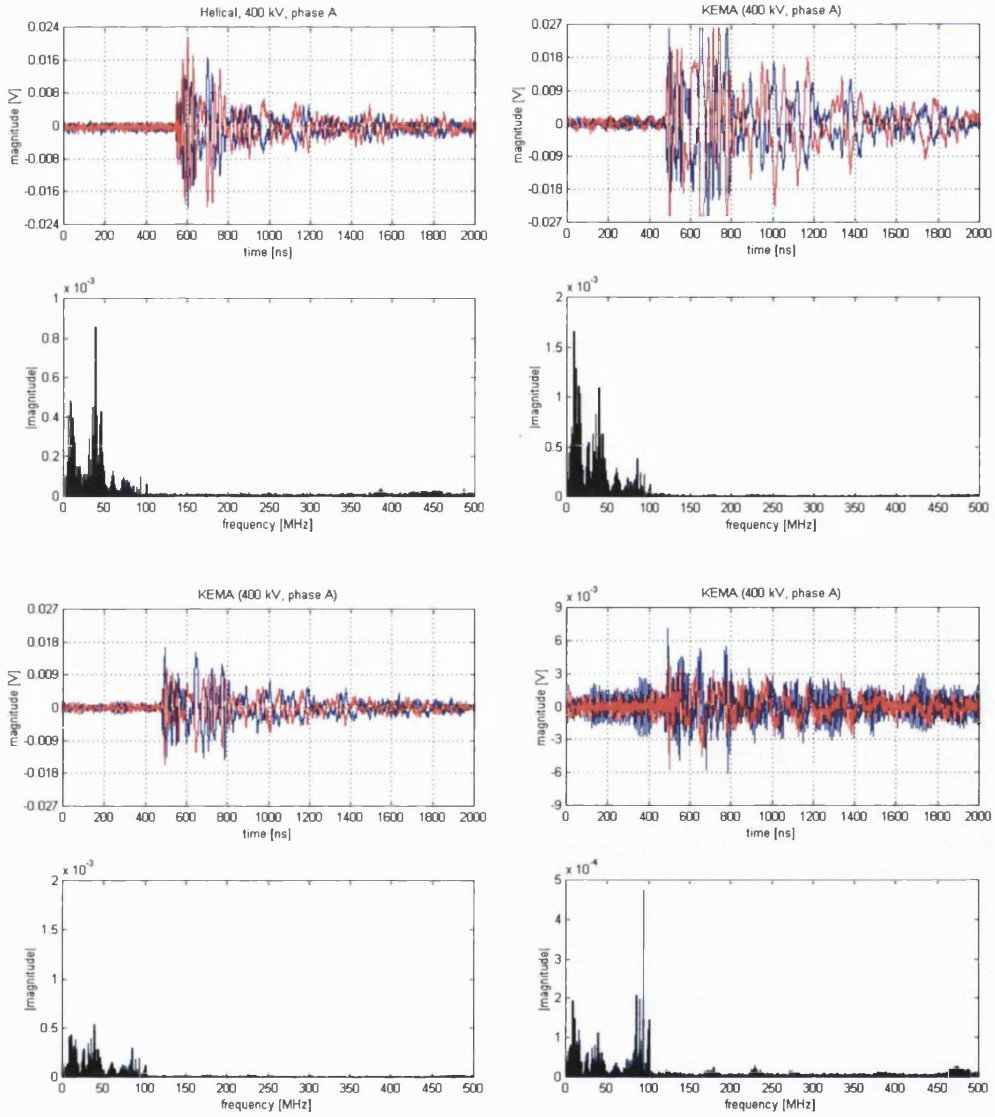


Figure 16.15: One group of signals measured by the helical antenna H3 and three groups of the KEMA probe signals

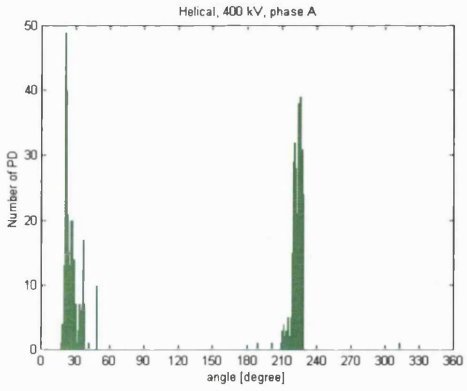


Figure 16.16: Distribution of the number of pulses regarding the voltage phase angle for H3

16.5. RESULTS OF THE MEASUREMENTS OBTAINED BY H4

Signals captured by the helical antenna H4, placed closed to phase C on 400 kV side only contain signals of one type, Figure 16.17, while signals obtained by the KEMA probe can be divided into four groups as shown in Figure 16.18.

Figure 16.19 presents the number of pulses with regard to the voltage phase angle. The distribution of the signals is between 15 - 50° and 130 - 140° and 200 - 3230°.

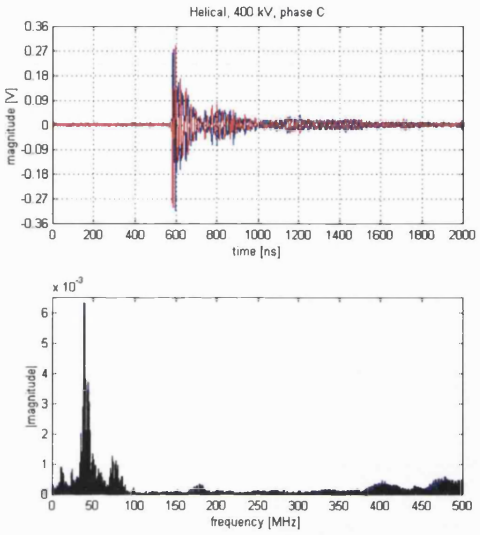


Figure 16.17: Signals measured by the helical antenna H4

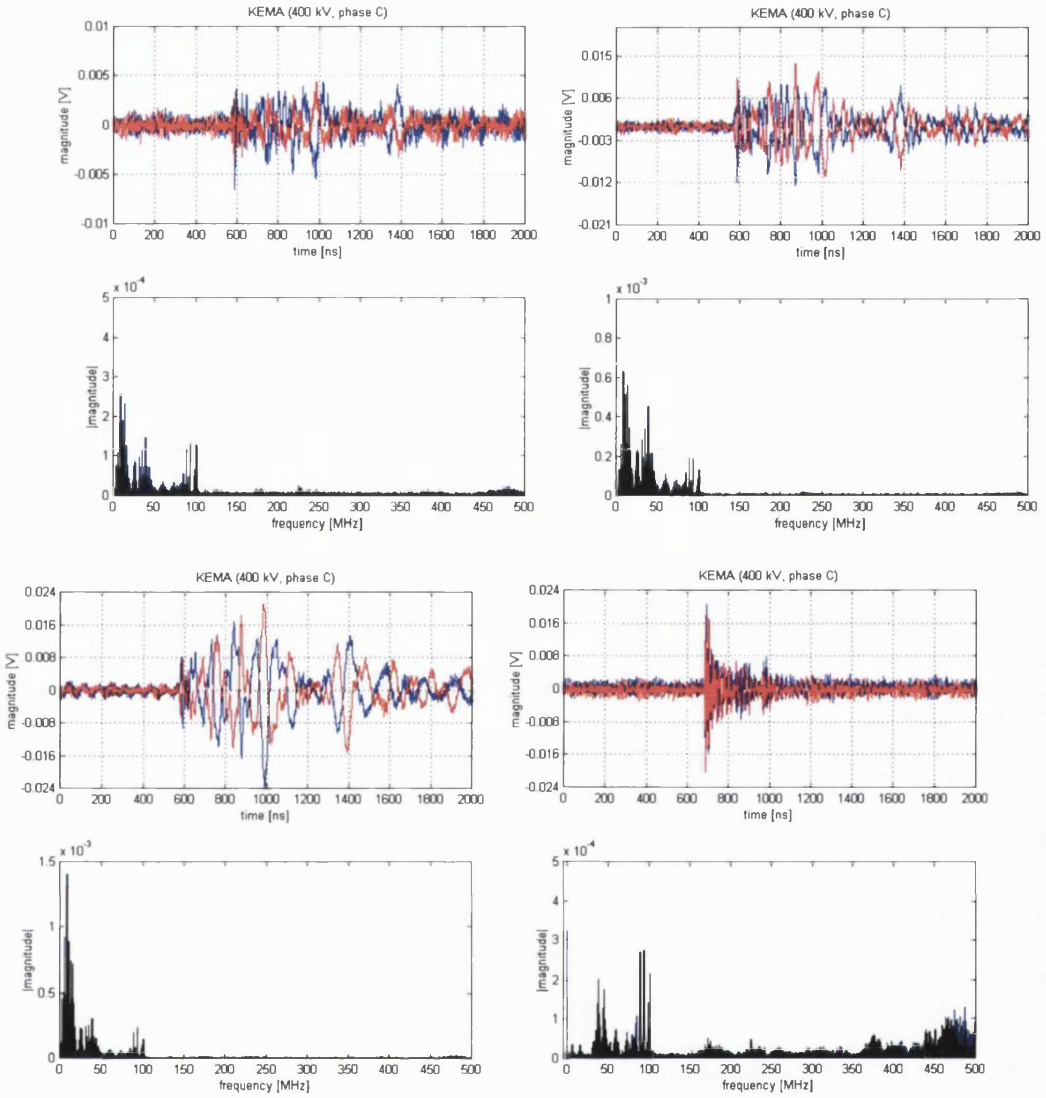


Figure 16.18: Four groups of signals measured by the KEMA probe for the same group of signals captured by the helical antenna (Figure 16.17)

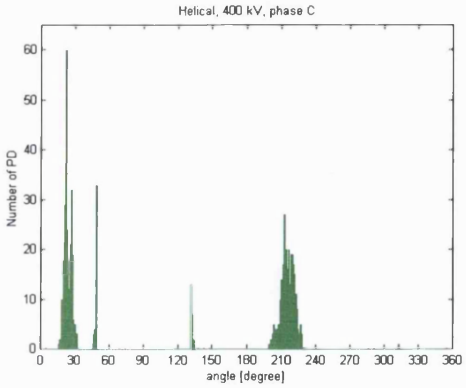


Figure 16.19: Distribution of the number of pulses regarding the voltage phase angle for H4

Related publications

T. Babnik, R.K. Aggarwal, P.J. Moore, Z.D. Wang: "Radio frequency measurement of different discharges", Power Tech Conference Proceedings, 2003 IEEE Bologna Volume 3, 23-26 June 2003, Page(s):5 pp. Vol.3.

T. Babnik, R.K. Aggarwal, P.J. Moore, P. Jarman: "Remote radiometric measurements of PD's occurring in power transformers", Proceedings of the XIIIth International Symposium on High Voltage Engineering, Delft, Netherlands, August 25-29, 2003. Rotterdam: Millpress, cop. 2003, pp. [1-4].

Paper accepted for presentation at 2003 IEEE Bologna PowerTech Conference, June 23-26, Bologna, Italy

Radio Frequency Measurement of Different Discharges

T. Babnik, R.K. Aggarwal, P.J. Moore and Z.D. Wang

***Abstract*—Discharges in power equipment may be a symptom of a fault in a component and in such cases their detection at an early stage is essential. This paper presents the detection and identification methods based on frequency and time domain measurements made with an antenna, oscilloscope and a laptop. The identification of oil discharge, floating particle discharge and air discharge by radio frequency interference measurements was studied at a High voltage laboratory. The findings of the research should serve as a basis for remote radiometric measurements of partial discharges occurring in power transformers.**

***Index Terms*—Partial discharge, oil discharge, floating discharge, air discharge, radio frequency measurement, signal detection.**

I. Introduction

Partial discharges (PD) are small discharges caused by strong and inhomogeneous electrical fields. The reason for such fields could, for example, be voids, bubbles or defects in an insulation material. Detection of PD is performed in order to ascertain the condition of the insulating material in high voltage elements, e.g. generators, transformers, gas isolated substations, cables, etc. Since PD usually occurs before complete breakdown, PD monitoring provides a warning to remove the power system element from service before catastrophic failure occurs.

For almost all systems in which PD is generated, the PD pulse originates with duration of at most a few tens (corona can go to 100 ns) nanoseconds. However, the signal, which is detectable outside the object, depends on the nature of the connection between the point of generation and the external world.

A variety of techniques are available to detect PD activity, the most common being acoustical [1, 2], optical [3] chemical [4], and electrical [5] techniques. Chemical methods are based on the analysis of dissolved gas (DGA) generated inside the transformer as a result of PD activity. Frequently performed DGA allows indications on the long-term behavior of the PD activity and therefore on the insulation condition. Information regarding actual PD occurrence requires acoustic or/and electric PD measurements. The focus of acoustic measurements is based on a PD location, whereas the electric measurements are orientated to an accurate determination of the apparent charge.

The use of remote radiometric measurements for identifying partial discharge behavior in high-voltage plant has not been investigated in any great detail in the past. This technique is potentially advantageous in field measurements, because it requires no galvanic connections to the testing object.

This paper presents a practical set-up for generating different discharges to investigate the level of electromagnetic radiation generated by discharges and their detection with wideband helical antenna. The main part is focused on the analysis of results present in time and frequency domain. The paper will conclude with summary of different characteristics and effects of electromagnetic radiation generated by different discharges.

II. PD types and their generation systems

For the purpose of this study, the following PDs were generating:

1. *Oil discharge*, produced with a point/pressboard/plate electrode arrangement with the distance between the point electrode and the pressboard about 5mm, Fig. 1. The pressboard is about 6mm thick. The arrangement is immersed in transformer oil during the test. The system is discharge-free up to the test voltage level of 40kV.

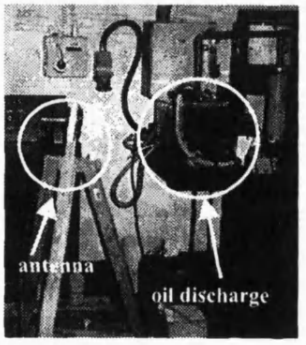


Fig. 1. Oil discharge source

This work was supported in part by the National Grid Company.

T. Babnik is with University of Ljubljana, Faculty of Electrical Engineering, Trzaska 25, SI-1000 Ljubljana, Slovenia (e-mail: tadeja.babnik@fe.uni-lj.si).

R.K. Aggarwal is with University of Bath, Electronic and Electrical Eng. Claverton Down, Bath, BA2 7AY, UK (e-mail: R.K.Aggarwal@bath.ac.uk).

P.J. Moore is with University of Bath, Electronic and Electrical Eng. Claverton Down, Bath, BA2 7AY, UK (e-mail: P.J.Moore@bath.ac.uk).

Z.D. Wang is with UMIST, Electrical Engineering & Electronics. Manchester M 60 1QD, UK (e-mail: zhongdong.wang@umist.ac.uk).

0-7803-7967-5/03/$17.00 ©2003 IEEE

2. *Floating discharge*; produced with a point/pressboard/ plate electrode arrangement, with two irregular balls made of aluminum foil floating in oil near the HV electrode, Fig. 2. These two aluminum balls are physically half immersed in oil and half floating above the oil. They are also electrically with a floating potential, which depends on how voltage is induced/coupled on to them. Only under lower voltage up to 3~4kV discharge can be produced.

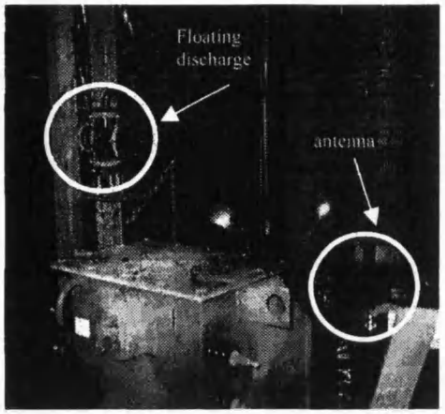


Fig. 2. Floating discharge source

3. The 11 kV transformer winding is used for generating *air discharges*, Fig. 3. The winding has 72 discs and each disc contains about 6 to 7 turns depends on their physical position.

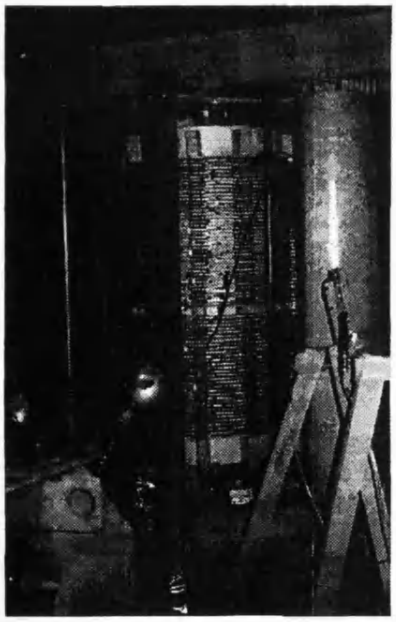


Fig. 3. Transformer winding used for generating discharge sources

It is a continuous disc type transformer winding. Discharge sources were connected to various disc positions to test how the position of discharge may affect the detection of the discharge signals.

III. Measurement Configuration

The electromagnetic radiation generated by different type of discharges was measured by a helical (frequency range 300-500 MHz) antenna, digital phosphor oscilloscope, and laptop, Fig. 4. In order to preserve the shape of the discharge pulse, the wide band measurements were performed. The oscilloscope enables acquiring data with frames. The so-called "FastFrame" is an acquisition mode that lets us capture many records in a larger record, and then view and measure each record individually. The setting of FastFrame was 100 frames of 5000 samples each. The measurement window was 2μs with $\Delta t = 4 \cdot 10^{-10}$ s.

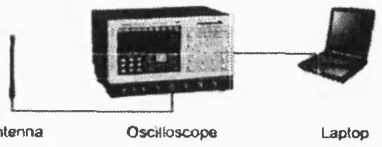


Fig. 4. The measurement configuration used for radio frequency interference measurements

IV. Measurement Results and Analysis

Measurements of discharges can be disturbed by noise originated from communication services and amplitude-modulated broadcasting stations. Therefore, the antenna captured the background noise before generating any type of discharges. Fig. 5 shows the background noise signal and its frequency spectrum. It can bee seen that the frequency spectrum of the background gives signals from FM radio (at 100 MHz), pagers (at 150 MHz) and DAB (at 220 MHz). The magnitude of the noise signal is smaller than 2 mV.

Fig. 6 shows a typical waveform of floating discharge along with frequency spectrum. Duration of captured signals is around 100 ns. Floating discharge contains high frequencies up to 600 MHz.

To influence of the distance between discharge source and antenna was also studied. The antenna was placed 219 cm, 230 cm and 326 cm away from the PD source. Fig. 7. compares waveforms measured at different locations from the source. Longer the distance is smaller is the magnitude of discharges. The frequency spectra of discharges are different in magnitude but have practically the same dominant frequency components.

Two examples of oil discharges are given in Fig. 8 and 9. Captured signals consist of single pulses (Fig. 8) or multiple pulses (Fig.9). The occurrence of the second discharge was approximately 600 ns after the first double pulse. Duration of pulses is around 200 ns. Single as well as multiple discharges in oil contain dominant frequencies below 300 MHz.

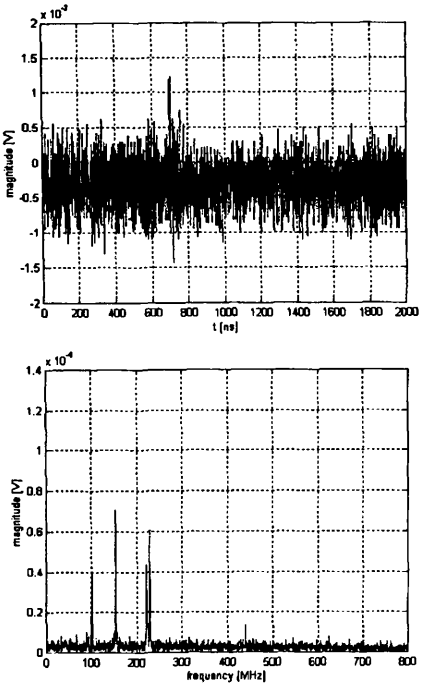


Fig. 5. The background noise of the signal and its frequency content

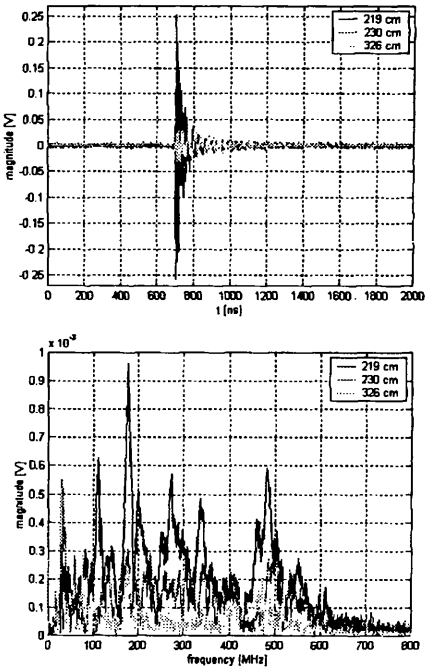


Fig. 7. PD waveforms measured at different locations and their frequency spectra

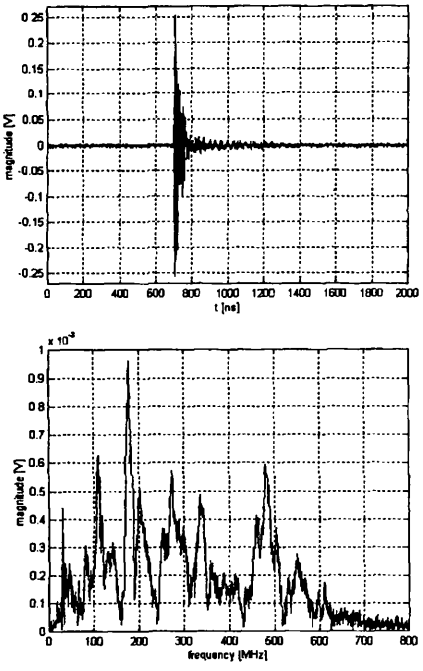


Fig. 6. PD signal caused by floating discharge and its frequency spectrum

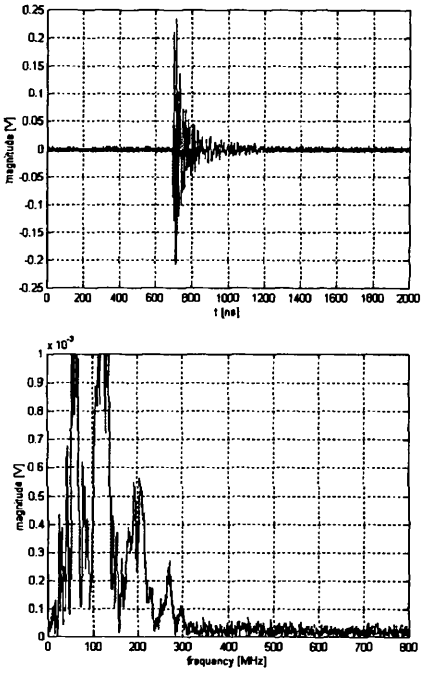


Fig. 8. Single pulse caused by oil discharge and its frequency spectrum

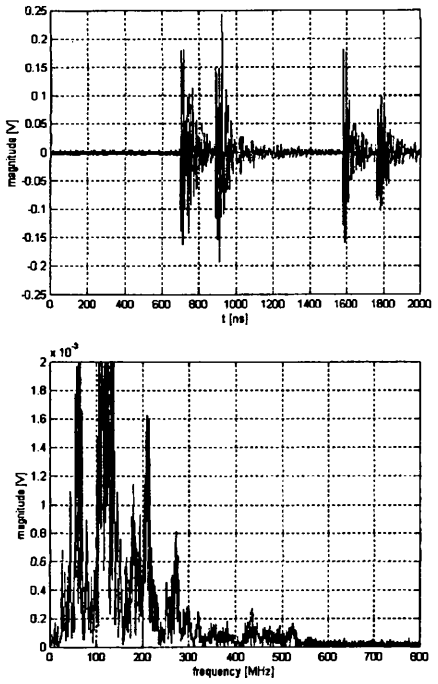


Fig. 9. Multiple pulses caused by oil discharge and frequency spectrum

Fig. 10 presents the typical example of air discharge and its frequency spectrum. Air discharge contains, in general, frequencies below 150 MHz. Duration of air discharges is longer then 800 ns. Generated air discharges were weak compare to discharges in oil or floating discharges; therefore the magnitude of air discharges is small compare to magnitude of discharges in oil or floating discharges. Similar wave shapes of signals are obtained for the disk position at the top and the bottom of transformer winding.

V. Conclusions

An oscilloscope equipped with the antenna and connected to a laptop proved to be a useful device in discharge type identification. According to the measurements it seems that different discharges can be evaluated by measuring the signals in time domain and frequency spectra of discharges. The difference between wave shapes and dominant frequencies can help with PD identification.

A simple goal of presented work is to identify possible variations and characteristics of discharge phenomena that provide mechanisms for future exploration of radio metric measurements for transformer PD measurements. Of course, measurements of PDs on-site will be disturbed by different noise. Therefore, it is important to create a database of different signals, which could serve for pattern recognition purpose.

Radio frequency measurements proved to be a useful tool in detecting and identifying different discharges. The presented technique is especially advantageous in field measurements, because it requires no galvanic connections to the equipment under examination.

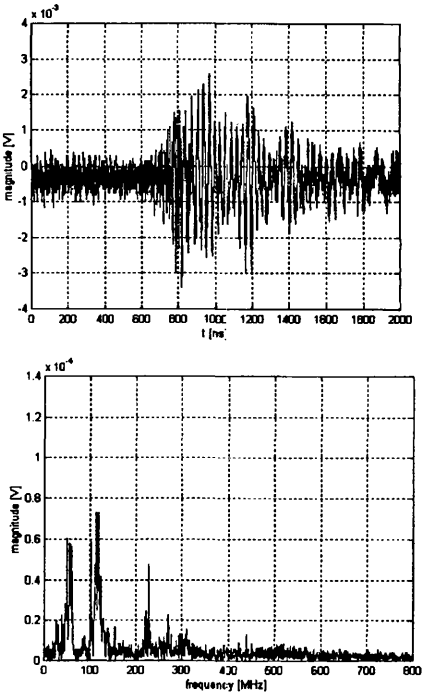


Fig.10. The shape of air discharge and its frequency spectrum

VI. Acknowledgement

The authors would like to thank to University of Bath, UMIST, the National Grid Company and the UK Radiocommunications Agency.

VII. References

[1] L.E. Lundgaard: "Partial discharge – Part XIII: Acoustic partial discharge detection – fundamental considerations", *IEEE Elect. Insul. Mag.*, 1992, 8, (4), pp 25-31.

[2] L.E. Lundgaard: "Partial discharge – Part XIV: Acoustic partial discharge detection – practical application", *IEEE Elect. Insul. Mag.*, 1992, 8, (5), pp 34-43.

[3] T.R. Blackburn, B.T. Phung, R.E. James: "Optical fibre sensor for partial discharge detection and location in high-voltage power transformer" Sixth International Conference on Dielectric Materials, Measurements and Applications (Conf. Publ. No.363). IEE. 1992, pp.33-6.

[4] M. Duval: "Dissolved gas analysis: it can save your transformer", *IEEE Elect. Insul. Mag.*, 1989, 8, (6), pp 22-27.

[5] A.J.M. Pemen, W.R Rutgers, T.J.M van Rijn, Y.H. Fu: "On-line partial discharge monitoring of HV components", Eleventh International Symposium on High Voltage Engineering, London, UK (Conf. Publ. No. 467), IEE. Part vol.5, 1999, pp. 136-139.

VIII. BIOGRAPHIES

Tadeja Babnik was born in Ljubljana, Slovenia, on April 14, 1969. She received Diploma Engineer and M.Sc. degrees in electrical engineering, from the University of Ljubljana, Slovenia, in 1993 and 1997, respectively. Since 1994 she has been working at Faculty of Electrical Engineering in Ljubljana. 2001 she joined Power and Energy System Group at the University of Bath, England for two years. Her current research interests are power system protection and control, and partial discharge monitoring in power transformers.

Raj K. Aggarwal received the B. Eng and PhD degrees from the University of Liverpool, England, in 1970 and 1973, respectively. He joined the Power and Energy System Group at the University of Bath, England, where he is now a Professor and Head of the Power and Energy System Group. His main research interests are power system modelling and application of digital-technology and AI to protection and control. He has published over 300 technical papers and co-authored four textbooks. Prof. Aggarwal is a Senior Member of the IEEE and Fellow of the IEE (U.K.).

Philip J. Moore was born in England. He received his BEng in Electrical Engineering from Imperial College London in 1984 and his PhD in Power System Protection from City University London in 1989. From 1984 to 1987 he was a Development Engineer at GEC Alsthom Protection and Control, formerly GEC Measurements. From 1987 to 1991 he was a lecturer in Electrical Engineering at City University. He joined the University of Bath in 1991 where he is presently a Senior Lecturer. Dr Moore is currently an investigator on a number of major research grants funded by SERC, GEC Alsthom and National Grid Company. Dr Moore has published numerous conference and journal papers and has been responsible for developing and running a highly successful short course on Numeric Power System Protection. Dr Moore has previously served on IEE Professional Group Committees P9 and P11 and organised the 27th Universities Power Engineering Conference held at the University of Bath, September 1992. Dr Moore's research interests include numeric protection, hierarchical protection systems, digital signal processing applied to power engineering, high voltage discharges, power system simulation and fault location.

Zhongdong Wang was born in Hebei, China, on 19 January 1969. She studied electrical engineering at Tsinghua University from 1986 until 1993, receiving her B. Sc. and M. Sc. degrees in 1991 and 1993, respectively. She was employed as a lecturer at Tsinghua University before coming to UMIST for doctoral studies. She completed her PhD at UMIST in July 1999 and subsequently joined UMIST as a Lecturer in February 2000. Her research interests are in high voltage engineering, partial discharge monitoring in transformers and cables, fast transient distribution in windings, ageing of insulation, material and diagnosis detection of winding displacements.

XIIIth International Symposium on High Voltage Engineering, Netherlands 2003, Smit (ed.)
© 2003 Millpress, Rotterdam, ISBN 90-77017-79-8

Remote radiometric measurements of PD's occurring in power transformers

T. Babnik[1], R.K. Aggarwal[1], P.J. Moore[1] and P. Jarman[2]
[1] University of Bath, Bath, United Kingdom
[2] National Grid Company, Coventry, United Kingdom

Abstract: Partial discharges in a power transformer may be a symptom of a developing fault; therefore their detection and evaluation at an early stage are essential. This paper presents partial discharge investigation carried out on the 400/275 kV autotransformer. The radiated electromagnetic waves caused by partial discharges are detected by a non-invasive measurement system consists of a wideband antenna, an oscilloscope and a laptop. Signals captured by a probe mounted on the oil valve at the bottom of the transformer tank serve as a reference for the signals obtained by radio frequency interference measurements. Attained data is presented in time and frequency domain.

1. Introduction

A power transformer is one of the main components in a power transmission network. Major faults in these transformers can cause extensive damage, which do not only interrupt electricity supply but also result in large revenue losses. Degradation in power transformer results in partial discharge activities. Hence, the detection of partial discharges (PD) forms a key element for many condition-monitoring strategies.

Partial discharges (PD) are small discharges caused by strong and inhomogeneous electrical fields. The reason for such fields could, for example, be voids, bubbles or defects in an insulation material. Detection of PD is performed in order to ascertain the condition of the insulating material in power transformers. Since PD usually occurs before complete breakdown, PD monitoring provides a warning to remove the power transformer from service before a catastrophic failure occurs.

A variety of techniques are available to detect PD activity, the most common being acoustical [1, 2], optical [3], chemical [4], and electrical [5] techniques. Chemical methods are based on the analysis of dissolved gas (DGA) generated inside the transformer as a result of PD activity. Frequently performed DGA allows indications on the long-term behaviour of the PD activity and therefore on the insulation condition. Information regarding actual PD occurrence requires acoustic or/and electric PD measurements. The focus of acoustic measurements is based on a PD location, whereas the electric measurements are orientated to an accurate determination of the apparent charge.

Use of remote radiometric measurements for identifying partial discharge behaviour in power transformers has not been investigated in any great detail in the past. This technique is potentially advantageous in field measurements, because it requires no galvanic connections to the transformer under test.

This paper is concerned with investigating the suitability of radio frequency interference (RFI) measurements for detection and identification of partial discharges occurring in 400/275 kV power transformer located on the NGC's transmission system. The main part of the paper is focused on the analysis of results presented in time and frequency domain. The paper will conclude with summary of different discharge shapes caused by different discharge sources.

2. Background

The autotransformer 400/275 kV, 500 MVA is connected to the NG transmission system. It has a history of producing gases in the oil indicative of partial discharge occurring within it. We try to investigate the RFI measurements for detection and identification of discharges occurring in power transformers. Well-established KEMA probe [5,6], for monitoring PDs was mounted on the oil valve at the bottom of the transformer tank. The signals from the probe serve as a reference signal for the measurements obtained by RFI measurements in order to separate signals caused by discharges outside the transformer and discharges occurring inside the transformer.

3. Measurement configuration

The PD measuring system consists of helical (frequency range 300-500 MHz) antenna, digital phosphor oscilloscope, and laptop, Figure 1.

Figure 2 shows a ground plan of the transformer and placement of antennas at four positions around the transformer. The letter K denotes the position of the KEMA probe.

Antennas on low voltage (LV) side were connected with the oscilloscope by 33m long semi rigid cables while antennas on high voltage (HV) side by 33m long satellite type cable.

The oscilloscope enables measurement of four quantities. The channel 1 captured data from one of the antennas (H1, H2, H3, or H4), channel 2 from the

KEMA probe and channels 3 and 4 served for capturing magnitude (V) and the slope (dV/dt) of the reference voltage.

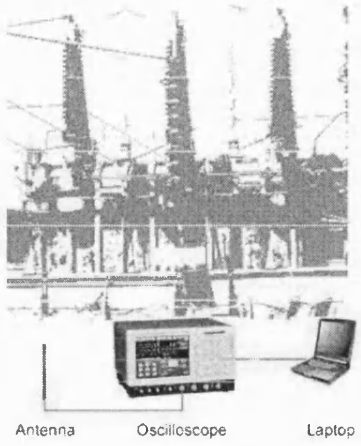


Figure 1: The RFI measurement configuration

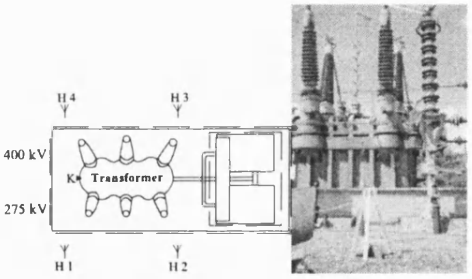


Figure 2: Ground plan of tested transformer and arrangement of helical antennas

In order to preserve the shape of the discharge pulse, the wide band measurements were performed. The oscilloscope enables acquiring data with frames. The setting of FastFrame was 100 frames of 5000 samples each. The measurement window was 2μs with $\Delta t = 4 \cdot 10^{-10}$ s.

4. Results and analysis

All signals presented in this paper were triggered by channel 1, i.e. by helical antenna placed close to the transformer. As mentioned before, signals obtained from KEMA probe serve as a reference for the signals obtained by antennas.

Signals detected by helical antennas and the probe contain different shapes and positions with regard to the reference voltage. In general, signals obtained at all locations can be divided into three to six representative groups. Each group usually contains two signals (mirror image) one occurring on positive half of the reference voltage and one occurring on the negative half of the reference voltage; both have nearly the same frequency spectrum.

Figures 3 and 4 show two out of 4 main groups of signals and their frequency spectra measured by H1, placed close to phase C on 275 kV side of the transformer (left) and signals detected by KEMA probe (right). The arrival time of the impulse at the H1 and KEMA probe is slightly different due to different lengths of connecting cables.

A waveform presented in Figure 3 shows the character of discharge originating from the transformer. If we compare the discharge spectra (blue) and spectra of background noise (red) of the signal captured by the probe we can see that discharge signal contains a broad range of frequencies between 200 and 800 MHz, indicative of PDs in oil [5]. The background noise is in general caused by 100 MHz FM radio signals and 500-600 MHz TV signals. The discharge signal obtained by heliacal antenna H1 also contains high frequencies in the range between 300 and 600 MHz.

Figure 4 shows typical waveform of discharge originating from outside of the transformer. Dominant spectral content of the signal captured by probe is below 200 MHz and has a character of corona discharge. The signal captured by helical antenna also contains dominant frequencies below 200 MHz.

The comparison of both waveforms of captured discharges shows that duration of corona discharges is much longer (400 ns) then duration of oil discharges. Also the magnitudes of discharges originating from the transformer are smaller compared to the magnitudes of corona discharges. Of course, the signal, which is captured outside the transformer, is corrupted because discharges are enclosed in the tank that reduces the direct radiation.

Since the reference voltage was also captured together with discharge pulses, data can be presented as phase resolved patterns. Figure 5 shows the number of pulses with regard to the voltage phase angle. Green bars illustrate distribution of all discharge groups together for measurements obtained by H1. Black bars present discharges for the selected group. Left graph gives the number of pulses with regard to the voltage reference angle for discharges originating from the transformer while right graph for discharges not originating from the transformer. The distribution of discharges occurring inside the transformer is between 25-50°, 130-140°, 190-220° and at 315°. Signals caused by discharges outside the transformer occurring between 0-30°, and 175-215°.

All other groups of signals obtained by H1 in general consist of frequencies below 100 MHz, typical for discharges in the air. Moreover, the signals

captured at locations H2, H3 and H4 contain frequencies below 100 MHz.

5. Conclusions

Generally, power transformers are very reliable elements, but when problems occur they are often difficult to diagnose and expensive to correct. Forced outage of a large transformer due to insulation breakdown may lead to catastrophic consequences for the transformer, associated equipment and cause environmental damage. Accurate assessment of the transformer insulation condition is therefore essential for safe and economic operation.

Remote radio frequency measurements proved to be a useful tool in detecting and identifying different discharges occurring in power transformers. The advantage of the presented method is that measurements can be performed when a transformer is in operation and it doesn't require any disconnection of the transformer from the power system. According to the measurements it seems that different discharges can be evaluated by measuring the signals in time domain and frequency spectra of discharges.

Based on the results of investigation following conclusions can be drawn. Helical antennas placed in the vicinity of a power transformer enable detection of multiple discharges. Correlation with KEMA probe signals allows separation of discharges occurring inside transformer from discharges occurring outside the transformer. At locations H2, H3 and H4 pulses with dominant spectral content below 100 MHz were obtained. Pulses with dominant spectral content above 200 MHz, indicative of PDs in oil, were detected only at H1 position close to phase C on 275 kV side.

For pattern recognition purpose it is important to create a database of different signals, therefore more tests should be performed also on transformers not having problems with partial discharges in order to obtain better knowledge about pulses occurring on transformers not containing discharges.

6. Acknowledgement

This work was supported in part by the National Grid Company. The authors would like to thank to University of Bath, the National Grid Company and the UK Radiocommunications Agency. The authors would also like to thank Iliana Portugues in view of her input regarding the FastFrame programming and point-on-wave calculation.

7. References

[1] L.E. Lundgaard: "Partial discharge – Part XIII: Acoustic partial discharge detection – fundamental considerations", *IEEE Elect. Insul. Mag.*, vol.8, no.4, pp 25-31, 1992.

[2] L.E. Lundgaard: "Partial discharge – Part XIV: Acoustic partial discharge detection – practical application", *IEEE Elect. Insul. Mag.*, vol.8, no.5, pp 34-43, 1992.

[3] T.R. Blackburn, B.T. Phung, R.E. James: "Optical fibre sensor for partial discharge detection and location in high-voltage power transformer" Sixth International Conference on Dielectric Materials, Measurements and Applications (Conf. Publ. No.363). IEE, pp.33-6. 1992

[4] M. Duval: "Dissolved gas analysis: it can save your transformer", *IEEE Elect. Insul. Mag.*, vol.8, no.6, pp 22-27, 1989.

[5] A.J.M. Pemen, W.R Rutgers, T.J.M van Rijn, Y.H. Fu: "On-line partial discharge monitoring of HV components", Eleventh International Symposium on High Voltage Engineering, London, UK (Conf. Publ. No. 467), IEE. Part vol.5, pp. 136-139, 1999

[6] W.R. Rutgers, P. van den Aardweg, A. Lapp, H.G. Kranz: "Transformer PD measurements: Field experience and automated defect identification", XIII Int. Conf. on gas discharges and their applications, Glasgow, 2000.

Authors' addresses
University of Bath, Electronic and Electrical Eng.
Claverton Down, Bath, BA2 7AY, UK
Email: eeptb@bath.ac.uk; R.K.Aggarwal@bath.ac.uk; P.J.Moore@bath.ac.uk

National Grid Company, Coventry, CV4 8JY,UK
Email: Paul.Jarman@uk.ngrid.com

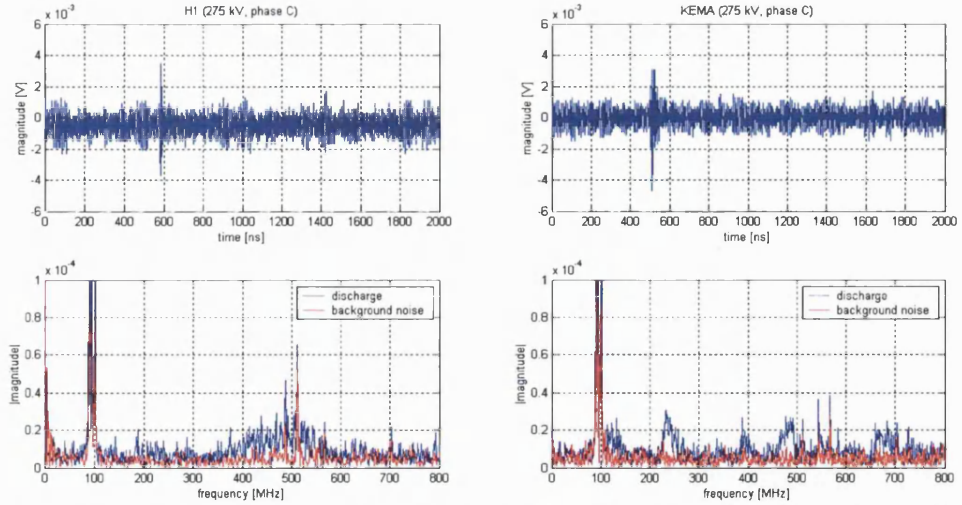


Figure 3: Typical waveform and spectra of PD signal originating from the transformer (captured by antenna on position H1–left, captured by KEMA probe – right)

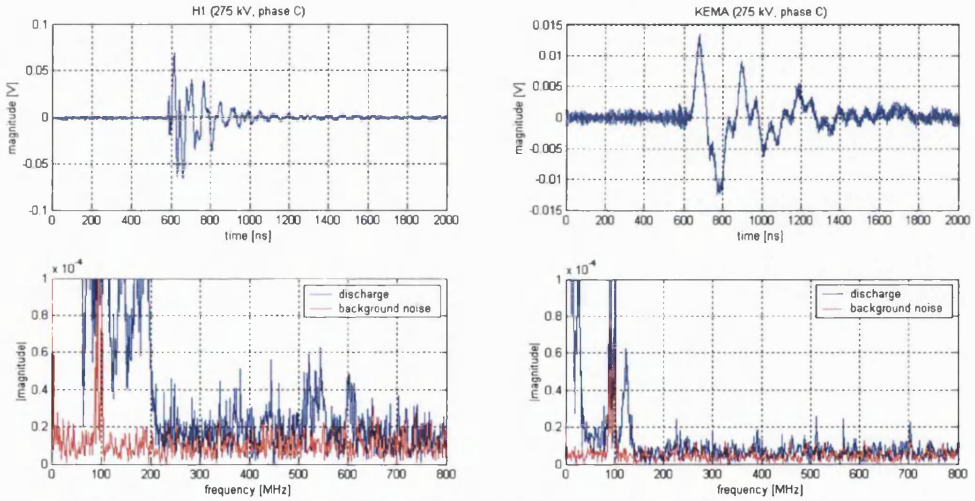


Figure 4: Typical waveform and spectra of PD signal caused by discharges outside the transformer (captured by antenna on position H1 – left, captured by KEMA probe – right)

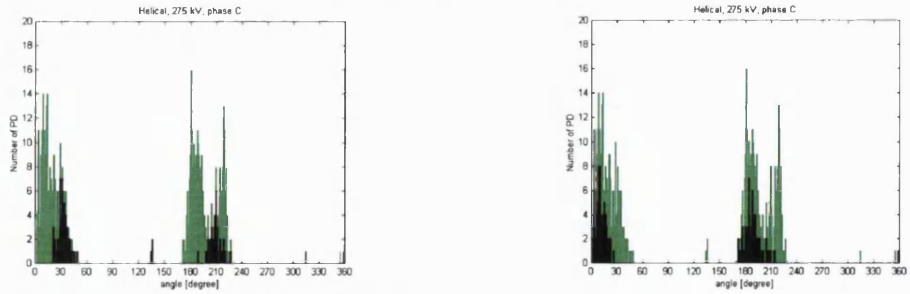


Figure 5: Distribution of the number of pulses regarding the voltage reference angle for discharges originating from the transformer (left) and for discharges not originating from the transformer (right)